The development and evaluation of an evidence-based approach to implementing outcome measurement in routine mental health services

Slade, Michael Dominic

The development and evaluation of an evidence-based approach to implementing outcome measurement in routine mental health services

Michael Dominic Slade

Institute of Psychiatry
King's College London

Submitted to the University of London

for the degree of

Doctor of Philosophy (PhD)

March 2004

Contents

Page

Title 1

Contents 2

List of Tables 6

List of Figures 7

Abbreviations 8

Glossary 10

Acknowledgements 13

Dedication 14

Declaration 15

Abstract 16

Chapter 1 – Introduction 17
1.1 Aims of the study 17
1.2 Outcome 17
1.3 Outcomes management 18
1.4 Measuring outcomes in mental health services 19
1.5 Scientific framework 21
1.6 Overview of the thesis 23

Chapter 2 – Outcome domains 25
2.1 Choosing outcome domains 25
2.2 Method of systematic review 26
2.2.1 Review quality 26
2.2.2 Search strategy 26
2.2.3 Data sources 28
2.2.4 Data abstraction 29
2.3 Results – outcome domain categories 31
2.4 Discussion of outcome domain categories 37

Chapter 3 – Routine assessment of outcome 39
3.1 Assessing outcomes in routine clinical practice 39
3.2 Current practice in routine use of outcomes 40
3.2.1 Current practice in the United States 40
3.2.2 Current practice in Australia 41
3.2.3 Current practice in England 42
3.2.4 Routine use of outcomes at different levels 43
3.3 Routine outcome assessment may directly benefit patients 44
3.3.1 Routine outcome assessment in psychotherapy 44
3.3.2 Routine outcome assessment in adult mental health services 46

3.3.2.1 Reflective practice 46
3.3.2.2 Types of problems 48
3.4 Results – principles for routine assessment of outcome 50
3.5 Internal validity of the systematic review 52
3.6 External validity of the systematic review 56
3.7 Discussion of principles to inform the routine assessment of outcome 56
3.8 Four-step approach to assessing outcomes routinely 57
3.8.1 Proposed Step 1: Develop the model 57
3.8.2 Proposed Step 2: Choose the outcome domain(s) 57
3.8.3 Proposed Step 3: Resolve the technical issues 59
3.8.4 Proposed Step 4: Choose the outcome measure(s) 61
3.8.4.1 Most outcome measures are not designed for routine use 62
3.8.4.2 Developing feasible measures 63
3.8.4.3 Available measures for routine clinical use 64

Chapter 4 – A model of routine outcome assessment 69
4.1 Outcomes research informing the model 69
4.1.1 Target outcome domains for improvement 69
4.1.2 Feedback 71
4.2 Implementation step 1: Develop the FOCUS Model 72
4.3 Implementation step 2: Choose the outcome domain(s) 75
4.4 Implementation step 3: Resolve the technical issues 76
4.5 Implementation step 4: Choose the outcome measure(s) 79

Chapter 5 - The relationship between needs and quality of life 83
5.1 From model to theory 83
5.2 Research investigating unmet need 83
5.3 Research investigating quality of life 87
5.4 Hypotheses for confirmatory re-analysis 90
5.5 Method of confirmatory re-analysis 91
5.5.1 Confirmatory re-analysis: Participants 91
5.5.2 Confirmatory re-analysis: Procedures 92
5.5.3 Confirmatory re-analysis: Measures 92
5.5.4 Confirmatory re-analysis: Methods of Analysis 92
5.6 Results of confirmatory re-analysis 93
5.6.1 Hypothesis (i): Cross-sectional association and isolation 95
5.6.2 Hypothesis (ii): Longitudinal association 97
5.7 Discussion of confirmatory re-analysis 99
5.8 Recent empirical data 101
5.9 The nature of routine outcome data 102
5.9.1 Reduced effect size due to shrinkage 102
5.9.2 Reduced effect size due to attenuation 103
5.9.3 Different relationships in routine data 103

Chapter 6 – Method 105
6.1 Aims 105
6.2 Hypotheses 105
6.3 Design 106
6.4 Management, monitoring and registration 106
6.5 Sample size 107
6.6 Sample setting and location 108
6.7 Eligibility criteria 108

6.8 Sample selection 109
6.9 Allocation 112
6.10 Measures 113
6.10.1 Postal questionnaire 114
6.10.2 Evaluation measures 116
6.11 Methods to enhance data quality 120
6.12 Pilot study 120
6.13 Procedure 121
6.13.1 Predisposing to change 122
6.13.2 Enabling and reinforcing change 122
6.13.3 Recruitment 123
6.14 Description of intervention and control 125
6.15 Follow-up 126
6.16 Protocol deviations 128
6.17 Analysis 128

Chapter 7 – Results 132
7.1 Baseline results 132
7.2 Intervention provision 134
7.3 Retention 135
7.4 Follow-up results 137
7.5 Testing of hypothesis 1 141
7.6 Testing of hypothesis 2 142
7.7 *A priori* sub-group analysis 1: Premorbid IQ 142
7.8 *A priori* sub-group analysis 2: Staff group 148
7.9 Evaluation of the FOCUS Model 149
7.10 Testing of hypothesis 3 153

Chapter 8 - Discussion 158
8.1 Summary of findings from the FOCUS RCT 158
8.2 Effectiveness of the intervention 158
8.2.1 Candidate reason 1: Low implementation 159
8.2.2 Candidate reason 2: Contamination 159
8.2.3 Candidate reason 3: Lack of power 160
8.2.4 Candidate reason 4: Sample heterogeneity 160
8.2.5 Candidate reason 5: Effectiveness 162
8.3 Feasibility of the intervention 166
8.4 Patient-rated unmet need and quality of life 167
8.5 Limitations of the FOCUS RCT 168
8.5.1 Limitation 1: Sample selection 168
8.5.2 Limitation 2: Standardisation of measures 169
8.5.3 Limitation 3: Data completeness 170
8.5.4 Limitation 4: Contamination 170
8.5.5 Limitation 5: Blinding 170
8.6 Future work 172
8.6.1 Future work1: Premorbid IQ 172
8.6.2 Future work 2: Amending the FOCUS Model 174
8.6.2.1 Increasing the predisposition to behaviour change 175
8.6.2.2 Enabling behaviour change 177
8.6.2.3 The amended FOCUS Model 178
8.6.3 Future work 3: Providing care based on unmet need 180
8.6.3.1 A needs – quality of life model 180

8.6.3.2 A new type of adult mental health service 184

Chapter 9 – Conclusions 185
9.1 Evaluation of a complex health intervention 185
9.2 Scientific implications 187
9.2.1 Implication 1: Scientific framework 187
9.2.2 Implication 2: The FOCUS Model 188
9.2.3 Implication 3: Methodology 189
9.2.4 Implication 4: Outcome measure 189
9.3 Clinical implications 190
9.3.1 Implication 5: Changing staff and patient behaviour 190
9.3.2 Implication 6: Premorbid IQ 190
9.3.3 Implication 7: Patient-rated unmet need 191
9.4 Policy implications 191
9.4.1 Implication 8: Providing care on the basis of need 191
9.4.2 Implication 9: Routine outcome assessment 192

References 194

Appendix 1: Forms used in the FOCUS RCT 214

Appendix 2: Published papers 258

List of Tables

Table	Title	Page
2.1	Search strategy for the systematic review	28
2.2	Electronic databases used for the systematic review	32
2.3	Emergent categories of outcome domains for use in mental health services	34
3.1	Proposed frameworks for routinely assessing outcomes in mental health services	51
3.2	Checklist of issues to consider once the outcome domain is decided	60
3.3	Comparison of outcome measures intended for routine clinical use	66
4.1	Feasibility of selected outcome measures	81
5.1	Studies investigating the association between unmet need and other variables	85
5.2	Studies investigating the association between quality of life and other variables	88
5.3	Clinical and social characteristics of all patients (n=265) and long-term patient sub-group (n=121) at baseline	94
5.4	Regression on quality of life for all patients (n=265), using staff and patient assessments of need separately and combined	96
5.5	Regression on follow-up quality of life for long-term patients (n=121), using baseline variables with and without baseline quality of life included	98
6.1	Sample selection by age and gender	110
6.2	Sample selection by Community Mental Health Team	111
6.3	Point of administration and source of data for all assessments	119
7.1	Clinical and social characteristics of patients (n=160)	132
7.2	Baseline measures for the FOCUS RCT (n=160)	133
7.3	Postal questionnaire distribution and return rates	134
7.4	Follow-up measures for the FOCUS RCT	138
7.5	Researcher rating of allocation (n=143)	140
7.6	Outcome results by premorbid IQ status	144
7.7	Regression on follow-up patient-rated unmet need for premorbid IQ top quarter of patients (n=34), with (Model 3) and without (Model 2) follow-up variables	147
7.8	Change in follow-up patient-rated unmet need, by profession	148
7.9	Change in follow-up quality of life, by profession	149
7.10	Intervention group staff (n=81) and patient (n=85) responses to Impact of Involvement questionnaire	150
7.11	Care plan for intervention (n=100) and control group (n=58) patients at baseline and follow-up	151
7.12	Intervention group responses to Impact of Involvement questionnaire from top-quarter IQ patients and their staff (n=23)	153
7.13	Regression analysis of change in quality of life (dependent variable) with sociodemographic variables and change in clinical variables as independent variables (n=123)	155

List of Figures

Figure	Title	Page
1.1	The MRC Framework for Design and Evaluation of Complex Interventions to Improve Health	21
3.1	Lewin's model of experiential learning	47
4.1	FOCUS Model of routine outcome assessment and feedback leading to improved outcome through its effect on staff	73
4.2	FOCUS Model of routine outcome assessment and feedback leading to improved outcome through its effect on patients	74
5.1	Graphical chain model of baseline and follow-up needs and quality of life variables (n=121)	99
6.1	Summary of intervention group and control group timings for individual patients	127
7.1	Trial flow diagram for patients in the FOCUS RCT	136
7.2	Distribution of premorbid Full-scale IQ scores	143
7.3	Quality of life and patient-rated unmet need, by time point	156
8.1	Evidence from the FOCUS RCT for the staff version of the FOCUS Model	164
8.2	Evidence from the FOCUS RCT for the patient version of the FOCUS Model	165
8.3	Summary of FOCUS Model	174
8.4	The amended FOCUS Model	179
8.5	Unmediated model of the relationship between a mental health intervention and quality of life	181
8.6	Mediated model of the relationship between a mental health intervention and quality of life	181
8.7	General model of the relationship between mental health interventions and quality of life	183
9.1	Relationship between the FOCUS Study and the MRC Framework	188

Abbreviations

The following abbreviations are used:

95% CI	95% Confidence Interval
ANCOVA	Analysis of Covariance
BASIS-32	Behaviour and Symptom Identification Scale – 32 item
BPRS	Brief Psychiatric Rating Scale
CAN	Camberwell Assessment of Need
CANE	Camberwell Assessment of Need for the Elderly
CANSAS	Camberwell Assessment of Need Short Appraisal Schedule
CANSAS-P	Camberwell Assessment of Need Short Appraisal Schedule – Patient
CANSAS-S	Camberwell Assessment of Need Short Appraisal Schedule – Staff
CCS	Current Clinical Summary
CINAHL	Cumulative Index of Nursing and Allied Health Literature
CMHT	Community Mental Health Team
CNS	Cardinal Needs Schedule
CONSORT	Consolidated Standards of Reporting Trials
CPA	Care Programme Approach
CRD	Centre for Reviews and Dissemination
CUES	Carers and Users Expectations of Services
DAS	Disability Assessment Schedule
DRG	Diagnostic Related Group
ENMESH	European Network for Mental Health Service Evaluation
FACE	Functional Assessment of Care Environment
FOCUS	Feedback of Outcome to Users and Staff
GAF	Global Assessment of Functioning
HAS	Helping Alliance Scale
HAS-P	Helping Alliance Scale – Patient
HAS-S	Helping Alliance Scale – Staff
HoNOS	Health of the Nation Outcome Scale
IBSS	International Bibliography of Social Sciences
IQ	Intelligence Quotient
ISRCTN	International Standard Randomised Controlled Trial Number
LIG	Local Implementation Group
LQL	Lancashire Quality of Life Profile
LSM	Lehmann's Summary Measure
LSP	Life Skills Profile
MANSA	Manchester Short Assessment of quality of life
MHI	Mental Health Inventory
MRC	Medical Research Council
MS	Mike Slade
NART	National Adult Reading Test
NFCAS	Needs for Care Assessment Schedule
NHS	National Health Service
NICE	National Institute for Clinical Excellence
NNT	Number Needed to Treat
ns	Not significant
PAS	Patient Administration System
PRiSM	Psychiatric Research in Service Measurement
Q-LES-Q	Quality of Life Enjoyment and Satisfaction Questionnaire
QOL	Quality of life

QOLI	Quality of Life Interview
QOLIMH	Quality of Life Index for Mental Health
QUOROM	Quality of Reporting of Meta-analyses
R&D	Research and Development
ReFeR	Research Findings electronic Register
RCT	Randomised Controlled Trial
RFS	Role Functioning Scale
s.d.	standard deviation
SF-20	Short-Form General Health Survey
SF-36	Medical Outcomes Study Short Form
SLDS	Satisfaction with Life Domains Scale
SVOP	South Verona Outcome Project
TAG	Threshold Assessment Grid
TSC	Trial Steering Committee
UK	United Kingdom
VSSS	Verona Service Satisfaction Scale

Glossary

Attenuation	A reduction in the reliability of an outcome measure due to measurement error
Automated problem-solving	The non-conscious pattern-matching of the current situation with a previously developed problem-solving template to active the appropriate procedural knowledge
Cognitive dissonance	A psychological phenomenon which refers to the discomfort felt at a discrepancy between what you already know or believe, and new information or interpretation
Content of care	The interventions which are provided for patients
Control group	The 59 patients in the FOCUS RCT who received treatment as usual
Evaluation measures	The outcome measures assessed at baseline and follow-up to investigate the impact of the intervention
Feasibility	The extent to which an outcome measure is suitable for use on a routine, sustainable and meaningful basis in typical clinical settings, when used in a specified manner and for a specified purpose
FOCUS Model	The testable model of routine outcome assessment, presented in Chapter 4
FOCUS RCT	The randomised controlled trial to evaluate the FOCUS Model, described in Chapters 6 and 7
FOCUS Study	The whole body of work presented in this thesis
Ill-defined problems	A type of problem characterised by inadequate information on the initial state of the problem, the goal state, the legal operators (things that can be done to solve the problem) and the operator restrictions (constraints on the application of operators)
Intervention group	The 101 patients in the FOCUS RCT who received the same care as the control group, and in addition both the patient and their member of staff (i) were asked to complete a postal questionnaire on a monthly basis, and (ii) were both provided with identical feedback by post at 3-monthly intervals

Mediator	A variable which impacts on the strength or direction of the outcome
Minimisation	A randomisation approach in which the chance of allocation to a particular treatment is adjusted to account for any existing imbalances in the baseline characteristics of the sample.
Moderator	A variable which predicts the outcome of receiving mental health care
MRC Framework	The Medical Research Council Framework for Design and Evaluation of Complex Interventions to Improve Health
Outcome	The effect on a patient's health status attributable to an intervention by a health professional or health service
Outcome domain	A conceptually distinct component of outcome, such as quality of life, symptomatology or satisfaction with care
Outcome measure	A specific questionnaire or other form of assessment which measures a specified outcome domain
Outcomes management	A healthcare technology involving widespread use of standards in selecting interventions, routine and systematic measurement of outcomes, aggregation of this outcome data on a massive scale, and analysis and dissemination of portions of the database to decision-makers
Patient-based outcome monitoring	The monitoring of a target outcome by the patient at each therapy session, with the results charted over time and shared with the therapist
Postal questionnaire	The questionnaire sent to staff or to patients every month as part of the intervention
Process of care	How the content of care is provided
Prognostic factor	A factor which is thought likely to correlate with the patient's subsequent response to treatment
Reflective practice	The conscious reflecting on experience to link with theory and hence plan future experimentation
Routine service	A non-research based clinical setting, with no special resources available for assessing outcome which would not be available in other comparable standard services

Routine outcome assessment	The ongoing measurement and use of outcome data to inform decisions about whether to continue, change or curtail treatment
Shrinkage	A statistical property whereby the fit of the regression model to the data from one study is likely to reduce when the same variables are measured in another study of the same population and data quality
Standard psychometric properties	Psychometric properties such as reliability, validity and sensitivity change which need to be demonstrated as adequate before the outcome measure can be described as standardised
Systematic review	A review in which bias has been minimised by the systematic identification, appraisal and synthesis of all relevant studies on a specific topic according to a predetermined and explicit method
Unmet need	Exists where the patient experiences a current and serious problem, irrespective of any help given
Well-defined problems	A type of problem characterised by adequate information on the initial state of the problem, the goal state, the legal operators (things that can be done to solve the problem) and the operator restrictions (constraints on the application of operators)

Acknowledgements

Supervision for this work was jointly and generously provided by Prof Elizabeth Kuipers and Prof Graham Thornicroft. Their light touch and wise words have both been appreciated.

Many colleagues in the Health Services Research Department at the Institute of Psychiatry made valuable contributions. Administrative support was provided by Maura Butcher, Karen Langridge, Kirstin McLellan, Michelle Wilkins and Roza Wojcik. Dr Morven Leese provided significant and meaningful statistical support, Mauricio Moreno provided data management expertise, and Dr Paul McCrone provided health economics input. The researchers who worked on the FOCUS randomised controlled trial were Dr Sharon Cahill, Matt Gillard, Lisa McNally, Dr Alberto Parabiaghi, Preethi Premkumar and Dr Giovanni Salvi.

Some of the theory presented in this thesis emerged from creative discussions with Prof Stefan Priebe, whose substantial intellectual contribution I am happy to acknowledge.

Other theoretical aspects were developed during visits to the Department of Medicine and Public Health at the University of Verona, Italy. The visits were generously hosted by Prof Mirella Ruggeri and Prof Michele Tansella. The data reported in Chapter 5 came from the South Verona Outcome Project, which was originally funded by a grant from Fondazione Cassa di Risparmio di Verona, Vicenza, Belluno e Ancona, Progetto Sanità 1996-1997 ("The Role of Social Factors on Onset, Course and Outcome of Mental Disorders. An Epidemiological Approach" to Professor Michele Tansella). Considerable assistance was received from several Italian colleagues: Prof Annibale Biggeri, Dr Giulia Bisoffi and Dr Chiara Bonetto gave statistical advice, Dr Antonio Lasalvia provided data management advice, and Annalisa Voccia gave valuable administrative help.

The FOCUS randomised controlled trial was supported by two committees. The Trial Steering Committee comprised Prof Tom Burns, Dr Rose McCabe, Prof Stefan Priebe and Dr Chris Watkins. The Local Implementation Group comprised Mandy Addison, Penny Beckett, Caroline Carlisle, Dr Tim Crayford, Eileen Evans, Pam Filson, Dr Caron Gaw, Tony Goss, Dr Frank Holloway, Dr Kathy Kavanagh, Michael Larkin, Rosa May, Richard Pacitti and Tanya Reid. Janet Green provided the raw data on the sample from the PAS.

Finally, 160 service users and most mental health professionals working in Croydon took part in the FOCUS Study. Some Croydon staff had several participating services users – Charles Bailey, Edward Burrows, Marie Chin, Carolyn Chorlton, Michelle Duggan, Evelyn Elliott, Derek Fowle, Owen Hunter, Louise Jeal, Karen Kaouech, Claire Knight, Pauline La Forge, Dr Faisa Mufti, Jim Mullan, David Nartey, Richard Oola, Farid Pokun, Dr Sathananthan and Iqbal Surfraz. My sincere thanks to all participants.

The study was funded by a Medical Research Council Patient-Oriented Clinician Scientist Fellowship (Ref G108/390).

For Charlotte

Declaration

Versions of some of the work described in this thesis have been published elsewhere.

The systematic review of outcome domains (Chapter 2):

Slade M (2002) *What outcomes to measure in routine mental health services, and how to assess them – a systematic review*, Australian & New Zealand Journal of Psychiatry, **36**, 743-753.

The review of implementation principles for routine outcome assessment (Chapter 3):

Slade M (2002) *The use of patient-level outcomes to inform treatment*, Epidemiologia e Psichiatria Sociale, **11**, 20-27.

The FOCUS Model (Chapter 4):

Slade M (2002) *Routine outcome assessment in mental health services*, Psychological Medicine, **32**, 1339-1344.

The retrospective re-analysis of existing data from South Verona (Chapter 5):

Slade M, Leese M, Ruggeri M, Kuipers E, Tansella M, Thornicroft G (2004) *Does meeting needs improve quality of life?*, Psychotherapy and Psychosomatics, **73**, 183-189.

All publications are contained in Appendix 2.

I hereby declare that this dissertation, submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy and entitled "The development and evaluation of an evidence-based approach to implementing outcome measurement in routine mental health services", represents my own work and has not been previously submitted to this or any other institution for any degree, diploma or other qualification.

M. Slade

March 2004

Abstract

This thesis develops and evaluates an approach to using outcome measures in adult mental health services. A systematic review identified outcome domains and implementation principles for routine outcome assessment, and informed the development of an evidence-based model.

The model was tested in the Feedback of Outcome to Users and Staff (FOCUS) randomised controlled trial (ISRCTN 16971059). 160 patients from eight adult mental health teams in Croydon (South London) were recruited (101 intervention, 59 control). The intervention was monthly collection from, and three-monthly feedback of outcome information to, patients and staff for seven months. The hypotheses were:

1. The intervention will lead to 1.0 fewer patient-rated unmet needs on the Camberwell Assessment of Need Short Appraisal Schedule (CANSAS).

2. The intervention will lead to an increase of 0.25 points in quality of life on the Manchester Short Assessment (MANSA).

3. Baseline level of patient-rated unmet need will predict follow-up level of quality of life

Hypotheses 1 and 2 were not confirmed. The intervention promoted reflection, but was ineffective because it did not lead to behaviour change. However, receiving the intervention was effective in reducing patient-rated unmet need for the sub-group of patients in the top quartile of premorbid IQ (B=-3.3, p=0.004).

Hypothesis 3 was partly confirmed. Baseline patient-rated unmet need was associated with follow-up quality of life (B=-0.17, p<0.001), but not when other baseline variables were included. Change from baseline to follow-up in patient-rated unmet need was associated with change in quality of life (B=-0.09, p<0.001), even when controlling for the influence of change in other variables (B=-0.07, p=0.001). Cross-sectional time series regression on the repeated measures indicated that absolute levels of (but not change in) quality of life were associated with absolute levels of patient-rated unmet need one month previously (coefficient=-0.04, p=0.002).

Chapter 1

Introduction

1.1 Aims of the study

This study has three types of aims: policy, clinical and scientific. There is a growing demand for increased use of outcome measures by adult mental health services in England (Department of Health, 2001). The first aim of the study is to inform policy-making in this area. The second aim is to investigate whether clinical benefits for patients arise from using an evidence-based approach to the routine use of outcome measures in adult mental health services. A substantial body of outcomes research now exists, which allows exploration of the relationships between different outcome measures. The third aim of this study is to advance scientific knowledge about these relationships.

1.2 Outcome

The primary purpose of mental health care is to improve mental health (Thornicroft & Tansella, 1999). Outcomes are the means by which improvement can be discerned, and so are of central importance to assessing the impact of mental health care. This study develops and evaluates an evidence-based approach to using outcome measures in adult mental health services in England. The purpose of this chapter is to define terms, identify key issues and conceptual frameworks, and give an outline of the remaining chapters.

What is an outcome? Outcome in health care relates to those aspects of life which are impacted upon by treatment. Various definitions of outcome have been proposed, sixteen of which are listed by the UK Clearing House on Health Outcomes (1996). An **outcome** will be defined as '*the effect on a patient's health status attributable to an intervention by a health professional or health service*' (Andrews, Peters & Teesson, 1994, p. 3). An **outcome domain** is a conceptually distinct component of outcome, such as quality of life, symptomatology or satisfaction with care. An **outcome measure** is a specific questionnaire or other form of assessment which measures a specified outcome domain.

1.3 Outcomes management

This study is based on a healthcare technology proposed by Paul Ellwood in 1988. He wrote about the use of outcomes in health care in general, without specific reference to mental health. His analysis of North American health services identified several major issues:

- patients were uninformed about the quality of their care, often being asked to assess quality in terms of waiting time, politeness of staff, and other process factors
- commissioners were sceptical of the true value for money offered by the services for which they were paying
- clinicians were frustrated by the conflict between clinical imperatives and the financial concerns of service commissioners
- health service managers were finding it increasingly difficult to make informed decisions about resource allocation within services, due to a lack of relevant and reliable data.

Ellwood proposed that the way forward was to develop a technology of **outcomes management**, which he defined as "*a technology of patient experience designed to help patients, payers, and providers make rational medical care-related choices based on better insight into the effect of these choices on the patient's life*" (Ellwood, 1988, p. 1551). Outcomes management has four components:

1. widespread use of standards and guidelines by clinicians in selecting interventions
2. routine and systematic measurement of the functioning and well-being of patients, along with disease-specific clinical outcomes, at appropriate time intervals
3. aggregation of this outcome data on a massive scale
4. analysis and dissemination of results from segments of the data base which are most appropriate to the concerns of each individual decision-maker.

Outcomes management is a long-term endeavour, which facilitates practice change through the provision of ongoing feedback. Ellwood suggested that the time was right to consider its implementation because of developments in information technology systems, reliable outcome measures, better clinical governance standards, powerful databases, and more sophisticated analyses.

A key element of this proposed solution is the ongoing measurement of important outcomes, coupled with the use of this information to plan and evaluate care *at the level of the patient*. Data can of course be used at other levels, such as planning services nationally, but the focus of this study is on patient-level collection and use of outcome data. This practice will be referred to as **routine outcome assessment**, and corresponds to components 2 and 4 of outcomes management. It is defined in more detail in Section 3.2.4. The culture-shift implicit within this approach is that routine outcome assessment becomes seen as an integral component of clinical care, rather than an administrative burden added on to the 'real' work of clinicians.

Routine outcome assessment is an important topic, for both ethical and scientific reasons. Ethically, it is important to ensure that the treatment being provided in routine services is of the highest quality, and this can only be done by monitoring its impact. Scientifically, although a fair amount is known from research studies about the efficacy of a range of treatments, far less is known about the effectiveness or cost-effectiveness of treatments when used in routine mental health services. Motivations for routine outcome assessment therefore include maximising quality of care, determining the effectiveness of a treatment against doing nothing or doing something else, comparing effectiveness with published standards, and examining outcomes for sub-groups of patients (Ellwood, 1988; Weiss, 1998).

This study relates to the use of routine outcome assessment in adult mental health services. Some specific issues arising in this service context will now be briefly reviewed.

1.4 Measuring outcomes in mental health services

A range of practical issues in the measurement of outcome have been identified. One proposal by Clifford (1998) is that measurement of outcome should involve consideration of six issues or 'M's:

- Multi-axial – outcome should be measured in relation to more than one outcome domain
- Multi-perspective – outcome should be investigated from more than one perspective. Possible perspectives include the patient, the clinician, the carer, and the tax-payer

- Multi-functional – the information collected using outcome measures should serve more than one purpose. Possible purposes include informing clinical decision-making, demonstrating the service offers value-for-money compared with other services, and informing resource allocation decisions
- Multidisciplinary – outcome measures should be acceptable and useable by all members of the mental health multidisciplinary team, who are likely to include care managers, occupational therapists, psychiatric nurses, psychiatrists and psychologists
- Multi-agency – the information should be useable as a means of communication between agencies, especially primary care, secondary mental health services, and social services
- Multi-site – outcome measures should be suitable for use and aggregation across sites, both to establish norms for different settings and to enable benchmarking comparisons to be made.

To these six issues, a further important consideration can be added:

- Level of evaluation – three levels of mental health service can be differentiated (Burns & Priebe, 1996). At the treatment level, specific interventions such as medication, graded exposure and social skills training are given. At the programme level, combinations of different treatment components, such as a community mental health team or an early onset service, are provided. Finally, the system level comprises all programmes for a defined target group in a given area. Outcomes can be considered at each level.

Assessing outcome specifically in mental health services also requires consideration of at least four conceptual issues:

1. The effect of the treatment may be to slow decline or to maintain the current level, so the score on the outcome measure itself may not improve (or may even get worse) despite best quality clinical care
2. Current scientific development for mental health treatments is limited, so the difference between best and average quality care may be difficult to detect. For example, the best available evidence in the United Kingdom indicates that clinical

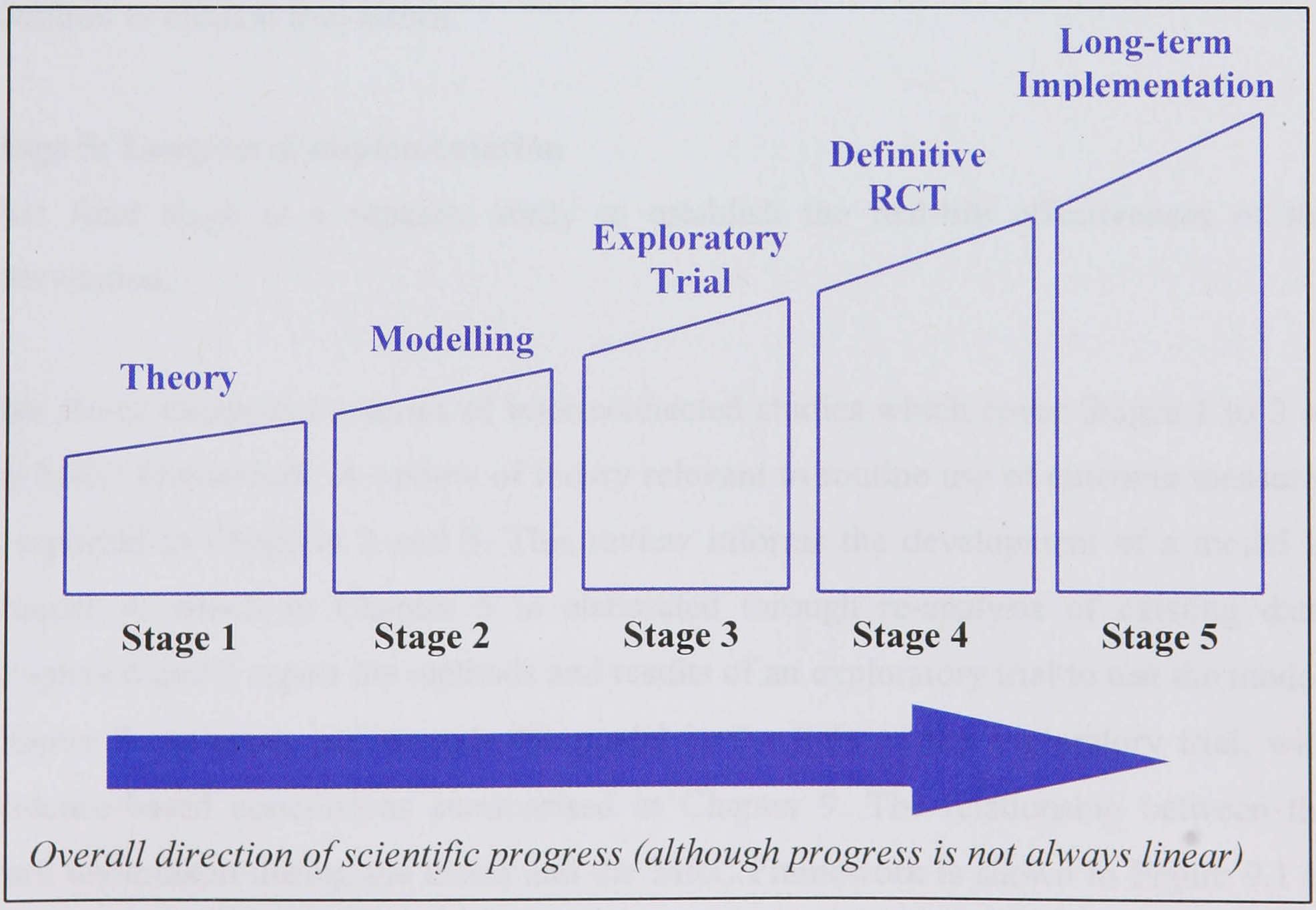
and social variables predict no more than 30% of the variance in an individual's quality of life (UK700 Group, 1999).

3. Different types of outcome are desynchronous (*e.g.* Drury, Birchwood, Cochrane et al, 1996), changing at different rates during an intervention
4. There may not even be agreement regarding what constitutes a positive change in outcome – the person who has fewer episodes of mania as a result of treatment may see this as a negative outcome.

1.5 Scientific framework

The scientific framework for this study is the Medical Research Council (MRC) Framework for Design and Evaluation of Complex Interventions to Improve Health (Medical Research Council, 2000), which has been published in summary form (Campbell, Fitzpatrick, Haines et al, 2000). It will be referred to here as the **MRC Framework**. This structural framework identifies five stages in designing and evaluating complex health interventions, which are shown in Figure 1.1.

Figure 1.1: The MRC Framework for Design and Evaluation of Complex Interventions to Improve Health



Each stage will now be described.

Stage 1: Theory development

This stage establishes the theoretical basis for the intervention, so as to identify in preliminary form the kind of intervention needed, and the type of study design which may be feasible.

Stage 2: Modelling

This stage involves developing an understanding of the intervention and its possible effects. This includes delineating an intervention's components and how they inter-relate, and identifying how the active components may relate to final outcomes.

Stage 3: Exploratory trial

This stage involves testing the evidence gathered during theory development and modelling stages. The goals may include establishing the feasibility of the intervention, experimenting with different components of the intervention, and providing sample size estimates for the next stage.

Stage 4: Definitive randomised controlled trial

This stage more formally evaluates the complex intervention, and requires more attention to clinical trial issues.

Stage 5: Long-term implementation

This final stage is a separate study to establish the real-life effectiveness of the intervention.

This thesis comprises a series of inter-connected studies which cover Stages 1 to 3 of the MRC Framework. A review of theory relevant to routine use of outcome measures is reported in Chapters 2 and 3. This review informs the development of a model in Chapter 4, which in Chapter 5 is elaborated through re-analysis of existing data. Chapters 6 and 7 report the methods and results of an exploratory trial to test the model. Chapter 8 evaluates and amends the model in the light of the exploratory trial, with evidence-based conclusions summarised in Chapter 9. The relationship between the work undertaken during the thesis and the MRC Framework is shown in Figure 9.1 in Chapter 9.

1.6 Overview of the thesis

The acronym Feedback of Outcome to Users and Staff (FOCUS) was chosen for this research programme, both because it describes the intervention and because it highlights that mental health service users and staff are both involved. For clarity, the model developed in Chapter 4 is called the **FOCUS Model**, the randomised controlled trial (RCT) to evaluate the FOCUS Model is called the **FOCUS RCT**, and the whole body of work presented in this thesis is called the **FOCUS Study**.

Chapters 2 and 3 use systematic review methodology to identify relevant theory. Chapter 2 presents a synthesis of the literature on proposals for outcome domains. The review identified sixteen separate proposals for groups of outcome domains to consider, which were synthesised into seven emergent categories of outcome domain. This synthesis provides a framework to bring order to the many differing proposals for which outcome domains to assess.

Moving from the general to the specific, Chapter 3 describes current practice in routine outcome assessment in three countries, and argues that preliminary evidence suggests it may be of direct benefit for patients. The literature is then systematically reviewed to identify principles which have emerged from previous attempts to introduce outcome assessment into routine practice. The results of the review inform the development of a four-step approach to implementing routine outcome assessment.

The four-step approach developed in Chapter 3 is then applied in Chapter 4. First, the theory reviewed in the previous two chapters is used to derive a testable model for routine outcome assessment – the FOCUS Model. Second, decisions about which outcome domains to assess are made, informed by the FOCUS Model. Third, technical issues which are relevant to the design of the FOCUS RCT and the implementation of the intervention are addressed. Finally, criteria are proposed for choosing outcome measures for use in routine services, and these criteria are applied to a group of outcome measures to identify specific measures for use in the FOCUS RCT.

In Chapter 5, theory is re-evaluated in the light of the FOCUS Model. A review of the literature on needs and quality of life indicates a potential causal relationship, with preliminary evidence suggesting that high levels of patient-rated unmet need may cause low levels of quality of life. This preliminary hypothesis is investigated through

retrospective re-analysis of an existing database – the South Verona Outcome Project. This re-analysis provides further preliminary evidence for a causal relationship between the two outcome domains, and allows the development of a testable hypothesis.

The FOCUS Model is then tested in an exploratory randomised controlled trial – the FOCUS RCT. Chapter 6 describes the method. The aims of the FOCUS RCT are to identify whether routine outcome assessment – the intervention, defined in Section 3.2 – leads to benefits for the patient, and to investigate whether the relationship between patient-rated unmet need and quality of life is preserved longitudinally. Three testable hypotheses are proposed, and the procedures involved in the FOCUS RCT are described.

Chapter 7 goes on to present the results of the FOCUS RCT. Descriptive data characterising the baseline sample, the feasibility of the intervention, and the follow-up data are reported. The three hypotheses are then tested. Two pre-planned sub-group analyses, relating to patients' premorbid IQ and the profession of staff participants, are also investigated. Evidence about the validity of the FOCUS Model is also presented.

Chapter 8 discusses the findings of the FOCUS RCT. The results in relation to each hypothesis are considered in the context of research published since the trial was planned. On the basis of the findings from the FOCUS RCT, an amended version of the FOCUS Model is proposed. A preliminary model for the influence of mental health interventions on quality of life is also outlined. Limitations of the study and potential future work are described.

Chapter 9 puts the FOCUS Study into context, and presents the main implications of the study for scientific research, clinical practice and mental health policy-making.

Chapter 2

Outcome domains

2.1 Choosing outcome domains

The choice of outcome domains to consider for use in adult mental health services is not self-evident. 'Hard' and often dichotomous clinical outcomes such as mortality and morbidity are used as end-points in many areas of medicine. Although mental health problems do reduce life-span (Brown, Inskip & Barraclough, 2000), the immediate goals of treatment normally relate to 'soft' outcomes such as reducing social disability or enhancing quality of life. Therefore morbidity and mortality are only partly appropriate for mental health services.

Publications describing mental health studies seldom describe why particular outcome measures (and, by implication, outcome domains) were chosen for use. This means at best that the basis for their consideration was implicit (*e.g.* based on expert view, or chosen to allow comparison with other studies), and at worst that there was no empirical basis for their inclusion.

As a first step towards a more evidence-based approach, it would be useful to have a framework for categorising outcome domains. This would inform decision-making about what to assess, allow a stronger justification for the inclusion of particular outcome domains, and provide a conceptual structure within which to investigate the relationship between different aspects of outcome.

The remainder of this chapter describes a systematic review of the literature from research relating to outcome assessment in mental health services for adults of working age. The review was undertaken with reference to **routine** adult mental health services. The term 'routine' means non-research based, *i.e.* with no special resources available for assessing outcome which would not be available in other comparable standard services. The two aims of the review were:

(i) to identify categories of outcome domains

(ii) to summarise principles for routine assessment of outcome.

Aim (i) will be met by reviewing proposals for outcome domains to assess, where the proposal is underpinned by some form of literature survey. The proposals will be synthesised into emergent categories of outcome domains. Aim (ii) will be met by reviewing proposed principles for implementing routine assessment of outcome.

2.2 Method of systematic review

A systematic review was undertaken by MS between September and October 2000. A systematic review is one in which bias has been minimised by the systematic identification, appraisal and synthesis of all relevant studies on a specific topic according to a predetermined and explicit method (Cook, Sackett & Spritzer, 1995). A meta-analysis was not appropriate because, although there is a wide evidence base on monitoring outcomes, the literature is not easily amenable to statistical aggregation.

2.2.1 Review quality

The quality of the review was maximised through the use of recognised search strategies (drawn from the York Centre for Reviews and Dissemination guidelines – www.york.ac.uk/inst/crd/search.htm), and by the use of the Quality of Reporting of Meta-analyses (QUOROM) guidelines to inform reporting of the results (Moher, Cook, Eastwood et al, 1999). Both of these sources are concerned with searches of clinical trials in medicine, and so were adapted for use when searching for research on this broader, conceptual theme. For example, guidance on quantitative data synthesis was not appropriate, and guidance on describing the information sources in detail informed the presentation of sources used (shown in Table 2.2).

2.2.2 Search strategy

The inconsistent use in the literature of key terms such as 'domain', 'routine', 'principle' and 'outcome' meant that it was not possible to identify a search strategy which differentiated between publications relating to Aim (i) (to identify categories of outcome domains) and Aim (ii) (to summarise principles for routine assessment of outcome), so the same search strategy was used for both reviews. For the same reason it was not possible to identify a search strategy which differentiated between publications relating to mental health research and to routine mental health services, so both were included for Aim (ii).

Identified articles were individually reviewed for core relevance, based on rater judgement. Some of the frameworks identified were described as models of "health status", "well-being" or "quality of life", but no search strategy was identified which allowed searching on any of these key-words with sufficient specificity, even though these concepts are conceptually distinct (Smith, Avis & Assmann, 1999). Similarly, no satisfactory search phrase for routine (as in "routine adult mental health services") could be found, so this aspect was incorporated where possible when reviewing abstracts (although often the distinction between research use and routine clinical use was not made explicit).

To maximise sensitivity, the search strategy was deliberately over-inclusive. Table 2.1 shows the search strategy used, in Medline format. The search was adapted for each database. For example, the International Bibliography of the Social Sciences (IBSS) database only allowed one search term, so "outcome" was used. No language restrictions were employed, and non-English articles were included where an abstract in English was available. Pre-publication studies and manuscripts which were in press were included.

Table 2.1: Search strategy for the systematic review

Search term	Meaning
1. outcome$.tw	*Any publication with "outcome" or "outcomes" in title or abstract plus…*
2. exp "Outcome and process assessment (health care)"/	*…any publication with this MeSH heading ORed with all its conceptually narrower terms…*
3. 1 or 2	*…gives anything to do with outcome.*
4. animal	
5. human	
6. 4 not (4 and 5)	*Studies only with animals…*
7. 3 not 6	*…will be excluded.*
8. (treatment or routine or review or assessment).tw	*At least one of these should feature.*
9. ("mental health" or psychiatr$).tw	*Only interested in mental health, either in abstract or title…*
10. exp psychiatry	*or in MeSH headings…*
11. exp mental health	
12. or/9-11	*…so combine.*
13. journal.pt	*Only interested in journals…*
14. multicenter study.pt	*…studies in more centre…*
15. practice guideline.pt	*…practice guidelines…*
16. review.pt	*…or reviews…*
17. or/13-16	*…so combine relevant sources.*
18. 7 and 8 and 12 and 17	
19. exp Therapeutics	*Treatment trials…*
20. 18 not 19	*…can be ignored*

2.2.3 Data sources

Findings from both unpublished studies and the 'grey' literature (*i.e.* reports which are not formally published, but available on the internet, in non-peer reviewed journals, and in other informal formats) were also considered, using three methods:

1. Researchers active in the field were consulted, and findings presented at the four European Network for Mental Health Service Evaluation (ENMESH) conferences held since 1995 were reviewed.

2. The world wide web was searched using Copernic 2000 (www.copernic.com), an internet site which collates the results from other internet search sites. Copernic 2000 was configured to search the world wide web using AltaVista, Direct Hit, EuroSeek, Excite, FAST Search, Google, HotBot, Infoseek, Lycos, Magellan, MSN Web Searcher, Netscape Netcenter, Open Directory Project, Web Crawler, and Yahoo!.
3. Two British registers of funded research were searched – the National Research Register which list research studies (completed and ongoing) which have relevance to the National Health Service (NHS), and the Research Findings electronic Register (ReFeR), which lists all centrally-funded NHS and UK Department of Health studies.

The main sources of information used are summarised in Table 2.2. Duplicates of all identified articles were removed using Reference Manager Professional Edition Version 9.5.

What time frame to use? A high-quality review was published by respected authorities (Andrews and colleagues) in 1994, so consideration was given to reviewing studies since 1993 (to allow for publication lag). This was intended to ensure the review was as systematic as possible, whilst remaining feasible. The Medline search for publications starting in 1993 yielded 1973 hits, whereas the search from 1966 yielded 2782 hits, indicating that 69% of articles in the last 35 years were published in the last 8 years. This accords with another study which found the annual rate of publications on outcome in mental health almost tripled between 1986 and 1996 (Trauer, 1998). Similarly, the CINAHL search from 1966-2000 yielded 275 hits, whereas searching from 1993-2000 yielded 255 hits. Interest in outcomes is relatively recent, suggesting that earlier references may be of less relevance. Therefore the search was restricted to publications in or since 1993. The reference list from all obtained articles were hand-searched, to identify key earlier references and primary sources.

2.2.4 Data abstraction

The titles of all publications identified in the search were scanned, and the abstract read where the title indicated possible relevance. Where the abstracts of matching publications indicated relevance, the full publication was obtained and read, following which a decision was made as to its inclusion.

The criteria used for inclusion are difficult to specify formally, since the identified publications were diverse in content and style. For Aim (i) (to identify categories of outcome domains), the inclusion criterion was that the proposal identified a range of (*i.e.* more than one) patient-level outcome domains for patients using adult mental health service, and that the proposal was underpinned by some form of literature review. Proposals relating to other areas of medicine were only included if the proposal was sufficiently generic to have relevance to mental health services, as rated by the reviewer. Exclusion criteria included a focus on choosing outcome measures (rather than domains), proposals relating to client groups other than adults of working age or outcome domains for programme (*e.g.* service uptake) or system levels (*e.g.* inter-agency working), reports of individual treatment trials, and having no literature review. After pilot searching with Medline, it was clear that none of these categories could be excluded with sufficient precision using a search strategy. The threshold for what constituted a literature review was also low, with no attempt made to judge the quality of review – any reference to a broader literature base beyond personal experience or expertise was included. What was specifically excluded was the use of expert opinion, whether it was implicit (where no attempt was made to justify the proposed framework) or explicit (such as a proposal made on the basis of "clinical experience").

For Aim (ii) (to summarise principles for routine assessment of outcome), the inclusion criterion was that a set of principles were proposed for assessing outcome in adult mental health services. The main reasons for exclusion were that proposals were too narrow (*e.g.* relating to minimising staff resistance to outcome measurement, relating to measuring outcome of psychotherapy, relating to outcome data solely for service funders, and – most commonly – relating to desirable psychometric properties of assessments) or too general (*e.g.* relating to measuring outcome in all medical settings).

Where more than one publication referred to the same piece of work, only the earliest was included, even where the apparently later one indicated that it was the first publication (*e.g.* Shern & Flynn, 1996; Smith, Manderscheid, Flynn et al, 1997). Similarly, where two publications varied only in relatively minor respects, such as the substitution of "Response to Care" with "Family and Informal Caregivers" between Clifford (1998) and Clifford (1999), the earlier publication was used. Where the date of 'publication' for grey literature was not clear, the date of the latest citation was used (*e.g.* Campbell, 1998).

For Aim (i) (to identify categories of outcome domains), the proposals were analysed to identify emergent categories. This analysis involved grouping outcome domains which overlapped between proposals, and using these groupings to identify the smallest number of emergent categories which both captured a coherent and conceptually-distinct set of outcome domains and covered the full range of proposals.

2.3 Results – outcome domain categories

The databases searched are shown in Table 2.2. The final column of Table 2.2 shows for each database the total number of publications matching either search criterion.

Table 2.2: Electronic databases used for the systematic review

Name of database	Brief description	Search engine	Web site	Search dates	Number found
Primary sources					
Medline	Biomedical research	Ovid version 7.8	www.biomed.niss.ac.uk	1993 – June 2000	1973
Cumulative Index of Nursing and Allied Health Literature (CINAHL)	Nursing research	Ovid version 7.8	www.biomed.niss.ac.uk	1993 – April 2000	255
PsycINFO	Psychology research	Ovid version 7.8	psycinfo.umds.ac.uk	1993 – April 2000	1941
International Bibliography of the Social Sciences (IBSS)	Social Sciences database	Ovid version 7.8	www.bids.ac.uk	1951 – June 2000	539
Research Findings electronic Register (ReFeR)	NHS R&D and Dept of Health Policy Research Programme	Unspecified	tap.ccta.gov.uk/doh/refr_web.nsf/basicsearch	1995 – Aug 2000	147
National Research Register 2000, Issue 3 (completed projects only)	National Health Service-relevant	Update software	www.update-software.com	Unspecified (64,000 records)	615
World Wide Web	"Grey" literature	Copernic 2000	www.copernic.com	August 2000	450
Secondary sources					
The Cochrane Methodology Register	Articles on the science of research synthesis	Update software	www.update-software.com	Unspecified	161
NHS Economic Evaluation Database	Economic evaluation of health interventions	NHS CRD	www.york.ac.uk/inst/crd	1994 – June 2000	248
Health Technology Assessment database	Information on health technology assessments	NHS CRD	www.york.ac.uk/inst/crd	1993 – Jan 2000	28

Sixteen specific proposals for outcome domains to measure which were based on some form of literature review were identified. Once all proposals were identified, they were synthesised into seven emergent categories, shown in Table 2.3. The rows of this table show the identified proposals ordered chronologically (to indicate the progress of ideas over time), and the columns show the emergent categories. The original terms for proposed outcome domains are used, listed under the emergent category or categories which best matched their definition (where given). The vertical lines indicate the emergent category or categories into which each specific outcome domain was placed – empty boxes in a row indicate that no outcome domain within the proposal matched the particular emergent category.

Table 2.3: Emergent categories of outcome domains for use in mental health services

	Study	Well-being	Cognition / Emotion	Behaviour	Physical Health	Interpersonal	Society	Services
1	Wenger, Mattson, Furberg et al, 1984	Perceptions of health status & well-being	Intellectual Emotional Symptoms	Daily routine	Symptoms - other illnesses	Social	Economic	
2	Bergner, 1985	Health perceptions General life satisfaction	Symptoms Emotional status Cognition	Functional status	Sleep and rest Energy and vitality	Role activities Social functioning		
3	Ware, 1989	Humanistic	Clinical			Rehabilitative	Public safety	
4	Hargreaves & Shumway, 1989	Mental health			Physical health	Social functioning Role functioning		
5	Rosenblatt & Attkisson, 1993	Life satisfaction and fulfillment	Clinical			Functional	Welfare and safety	
6	Ruggeri & Tansella, 1995	Quality of life Needs for care	Psychopathology			Social functioning and support	Burden of relatives	Satisfaction
7	Sederer, Hermann & Dickey, 1995		Symptom			Functional		Satisfaction with treatment
8	Cook & Jonikas, 1996	Quality of life				Vocational Educational Residential		Hospitalization Consumer satisfaction
9	McGlynn, 1996	Quality of life	Clinical			Functional		Adverse events Satisfaction with medical care
10	Schlosser, 1996	Life satisfaction Life direction	Emotional Mental		Physical		Social	

	Study	Well-being	Cognition / Emotion	Behaviour	Physical Health	Interpersonal	Society	Services
11	Campbell, 1998	Well-being Personhood	Recovery					Self-help Empowerment Iatrogenic effects Satisfaction and Dissatisfaction
12	Clifford, 1998		Psychological	Activities of daily living	Physical wellbeing	Interpersonal relationships	Social circumstances	Response to care
13	Fitzpatrick, Davey, Buxton et al, 1998	Psychological well-being Personal constructs	Cognitive functioning		Physical function	Social well-being Role activities		Satisfaction with care
14	Jennings, Staggers & Brosch, 1999	Health status / health-related quality of life Patient knowledge	Diagnosis Psychological function Symptom management	Behavioral Activities of daily living	Comfort / discomfort Physical function Mobility Disability	Social function		Patient satisfaction Appropriateness of treatment Sentinel events Technical proficiency
15	Thornicroft & Tansella, 1999	Quality of life	Symptom severity	Disabilities Needs			Impact of caring	Satisfaction with services
16	Thornicroft & Tansella, 2000	Quality of life		Needs Global functioning		Social disabilities	Carer burden	Quality of care Satisfaction

All literature reviews used as justification for the proposed outcome domains were non-systematic – no systematic review was identified. Study 14 stood out as being underpinned by a sound (though not systematic) literature review, and including aspects such as sentinel events (undesirable outcomes of a magnitude to always warrant a detailed investigation of a clinician's actions) and technical proficiency (of the clinician), which featured prominently in the general medical but not the mental health literature. (Technical proficiency is regarded as a process rather than an outcome in most mental health-focussed literature.)

The institutional affiliations of the authors of studies 1 to 5, 7 to 11, and 14 were North American, of study 6 were Italian, of studies 12 and 13 were English, and of studies 15 and 16 were English and Italian.

Studies 1 and 2 described components of health-related quality of life, studies 3, 10 and 13 described components of health status, and the remainder described treatment outcome domains. Study 1 related to a cardiovascular patient group, studies 2, 3 and 14 to general medicine, 10 to psychotherapy, 13 to clinical trials, and the remainder to patients of mental health services.

Most proposals defined the meaning of the outcome domain. For example, Ware (1989) defined 'Mental health' as both behavioural dysfunction and the frequency and intensity of symptoms of psychological distress and feelings of psychological well-being. 'Physical health' referred to limitations in performance, ability to perform daily self care, or undertake a range of physical activities. 'Social functioning' referred to both social contacts and social ties or resources. 'Role functioning' referred to performance of role activities such as employment, school and housework.

Hargreaves and Shumway (1989) stated that Humanistic goals are to maximise patient's and family members' sense of well-being and personal fulfillment. Clinical goals are to improve or cure an illness or disorder, reducing or eliminating its signs and symptoms. Rehabilitative goals are to restore or improve social and vocational functioning. Public Safety goals are to prevent injury whether from assaultive or self-destructive behaviours that arise out of illness, or from destructive side-effects of the services themselves.

Finally, Campbell (1998) described Well-being as linked to the protection of a person's basic human freedoms, safety and privacy. Personhood was a recognition of common humanity and a tolerance for individual differences. Self-help included both self-help groups and provision of specific services by consumers. Recovery was the optimisation of a consumer's life and the minimisation of their illness with appropriate, relevant and continuous flexible services and supports collaboratively developed and chosen. Empowerment involved the help receiver having direct control over the help and there being reciprocity between help givers and receivers. Iatrogenic effects and negative outcomes were undesired consequences from or side effects of receiving certain public mental health services or treatments. Satisfaction and dissatisfaction (distinct and so both important to measure) related to the consumer's view of services received and the results of the treatment.

The results for Aim (ii) (to summarise principles for routine assessment of outcome) are presented in Section 3.4.

2.4 Discussion of outcome domain categories

This study is the first systematic review of proposals for outcome domains for adults receiving care from mental health services. Synthesising previous work led to the emergence of seven categories of outcome domains: *Well-being*, *Cognition / Emotion, Behaviour*, *Physical Health*, *Interpersonal*, *Society* and *Services*. The internal and external validity of the review, methodological limitations and future work are discussed in Chapter 3, since some aspects relate to the results of the review of principles for outcome assessment.

Each emergent category can be considered independently. *Well-being* (*e.g.* life satisfaction) relates to the patient's perceptions about their life (not about services), and by definition can only be assessed by the patient. The next three categories relate to the patient in isolation from their context – their *Cognition / Emotion* (a single category spanning a range of intrapsychic aspects, such as symptom severity), their *Behaviour* (*e.g.* activities of daily living) and their *Physical Health* (*e.g.* mobility). For all three of these the clinician and the patient may have their own assessment, and their assessments may differ. The *Interpersonal* category (*e.g.* social functioning) refers to aspects of the patient in relationship to others, both in individual social interactions and in performance of social roles. The *Society* category describes aspects of a patient's mental

health problems which may impinge on wider society, both at the individual level (*e.g.* impact of caring), and the macro-level of costs (*e.g.* welfare benefits, public safety). Finally, the *Services* category includes both positive and negative aspects of receiving mental health care (*e.g.* satisfaction with care).

The order in which these categories are presented is intended to indicate a spectrum, with solely patient-defined Well-being at one end, through categories on which both patient and staff (and, for example, carers) may have their perspectives, to Society at the far end, for which the patient perspective may not be central. The Services category, relating to the experience of care, is distinct.

As a category, *Society* is the most diverse, encompassing both the micro-level impact of carer burden and macro-level effects, such as public safety. Notable by its absence was any proposal for outcome domains which elaborated these macro-level aspects. This might include the political, cultural and social dimensions of mental health, such as health beliefs, health seeking behaviour, stigma, and public models of mental disorder. The recent raising of concerns about whether institutional racism exists in mental health services in the NHS (Independent Inquiry, 2003) may lead to an increased focus on outcomes at the Society level. Future work might involve the disaggregation of this category into aspects concerning local impact (specifically, carer burden) and aspects concerning the wider societal impact of and perceptions about mental ill-health. Current conceptualisations, to which this review is limited, do not allow this separation, so a combined category has been used.

The emergent outcome domain categories offer a conceptual structure for service managers and clinicians who want to assess the impact of care on people using routine adult mental health services. Routine use of outcomes is discussed further in Chapter 3.

Chapter 3

Routine assessment of outcome

3.1 Assessing outcomes in routine clinical practice

The purpose of this chapter is to identify and apply theory relevant to the *routine* (as defined in Section 2.1) use of outcomes. This will involve first describing why the outcome assessment approach used in efficacy studies is inappropriate for routine use, and then in Section 3.2 surveying current practice in three countries. Evidence will then be presented in Section 3.3 which indicates that routine use of outcomes may have benefits at the patient level. Finally, the implementation principles identified in the systematic review described in Chapter 2 will be outlined in Section 3.4.

Choosing which outcome domains to assess is not simple, and nor is choosing the method of assessment or which informant perspectives to consider. In research studies, it is common for a wide range of outcomes to be measured from multiple perspectives (Biggeri, Bisoffi, Rucci et al, 2001; Clifford, 1998). For example, the PRiSM Psychosis Study evaluated two models of community care for people with psychotic diagnoses (Thornicroft, Strathdee, Phelan et al, 1998). The outcome domains assessed by interviewing the patient were symptomatology, needs, quality of life, services being received (to allow economic analysis), social network and satisfaction with care. The outcome domains assessed by interviewing staff were global level of functioning, needs and social behaviour. The outcome domains assessed by interviewing carers were their experience of care-giving and their own symptomatology. All interviews were conducted by highly-trained researchers. This is standard practice in most research and evaluation studies, which take place in '*research contexts where specifically funded and trained external raters parachute into routine clinical settings in order to guarantee the validity and reliability of study measures*' (Harrison & Eaton, 1999, p.187).

The research approach of using several outcome measures for each of several perspectives cannot be directly transferred to routine use for at least four reasons:

1. It requires the use of resources (*e.g.* interviewer time) which, whilst possible in research studies, are unlikely to be available in routine services. Monitoring even a small number of outcome domains in routine practice is time-consuming – Marks (1998) estimates an extra 10% of the clinician's time is involved.

2. It entails duplication of effort, when two outcome measures co-vary to the extent that one is a fair proxy for another
3. It can also be wasteful of effort in other ways, either when data are collected but not analysed, or when data are collected and analysed but do not inform future care planning or service development. It may be acceptable to absorb the adverse effects of duplication and waste of effort in research programmes, but in already over-stretched routine mental health services this is less defensible.
4. The outcome measures themselves may have limited *feasibility* (defined in more detail in Section 3.8.4.2), *i.e.* not be suitable for routine clinical use.

3.2 Current practice in routine use of outcomes

To inform decision-making about using outcomes routinely, current practice in assessing outcomes in routine adult mental health services will be reviewed. Examples from the United States, Australia and England will be considered.

3.2.1 Current practice in the United States

In the United States, the cost of mental health care as a proportion of total health expenditure rose from 3-4% in the early 1980s to nearly 25% in the early 1990s (Lyons, Howard, O'Mahoney et al, 1997). This resulted in the rapid introduction of managed care in the 1980s (Dickey & Azeni, 1992). Initially envisaged as a cost-containment procedure and driven to some extent by pressure from health insurance companies, the approach was based on the identification of Diagnostic Related Groups (DRGs), with each DRG being prospectively associated with access and 'benefit levels' (*i.e.* limitations on the care to be provided). The DRG system failed to result in equity and a fair matching of need with resources. This led to the emergence of 'carve-out' managed mental health care firms, which were entirely separately funded from the rest of the health care system in the mid 1980s. These were the precursor of the current arrangements involving managed care and health maintenance organisations. The net result of these changes was to make mental health care expenditure highly visible, and subject to market forces (since in general services were provided by for-profit organisations). This led to the DRG approach being superseded by an emphasis on outcomes rather than diagnosis, since diagnosis did not predict service use (McCrone & Strathdee, 1994).

Two main approaches have been introduced to attempt to assess the outcome of care offered – report cards and road maps. Report cards summarise the performance of the service at the programme or individual provider level, and are prepared on a regular basis (Lyons et al, 1997). Road maps attempt to link performance data from report cards with organisational functioning, to give information on the relationships between variations in processes and outcomes (White & Lyons, 1994). The intention is that outcomes achieved for patients should become the currency for evaluating services.

More generally in North America, there has been increased interest in outcome-based evaluation of social programmes (Schalock, 2000). The intention of this approach is to maximise quality: the link between process and outcome. In other words, the assumption is made that doing the right activity (process) to an adequate level of competence (quality) will produce outcomes of a specified level. If outcomes are not meeting the agreed level, then either the wrong treatment is being provided (in which case the process needs to be amended) or the right treatment is being offered with an inadequate level of competence (in which case quality needs to be improved). Several writers have used this deterministic approach to develop models for assessing and improving quality in mental health care (*e.g.* Dickey & Sederer, 2001; IsHak, Burt & Sederer, 2002). Within Ellwood's framework, the emphasis in these recent developments has been on the provision of information to two specific groups: those paying for and those managing the services.

3.2.2 Current practice in Australia

In Australia, the process of consumer involvement in mental health services is more advanced, and substantial efforts have been directed towards considering how outcomes should be monitored at the level of the individual patient. A seminal report by Andrews and colleagues (1994) identified candidate outcome measures for use in Australian mental health services. These measures were field-tested by an independent research team, resulting in specific recommendations for services (Stedman, Yellowlees, Drake et al, 2000). As a result, the staff-rated Health of the Nation Outcome Scale (HoNOS) (Wing, Beevor, Curtis et al, 1998) is now completed in some States for all patients using adult mental health services (Trauer, Callaly, Hantz et al, 1999). The emphasis has therefore been on the provision of information to service providers, mainly for informing local service developments.

3.2.3 Current practice in England

In England, as in the United States, cost has been a driver of change. Mental health services in England are increasingly expensive – in 1997/8 English Health Authorities spent £2,930million on mental health services (Bindman, Glover, Goldberg et al, 2000), and by 2002/3 this had risen to £5,800million (12% of the total expenditure of £48,655million) (Table E2 of Health and Personal Social Services Statistics, accessed from www.doh.gov.uk/HPSSS in December 2003). These figure exclude the substantial social service spend on mental health, as well as housing benefit and primary care mental health spending. It may therefore be unsurprising that in the 1990s, data collection in English mental health services was primarily, if not exclusively, driven by the demands of commissioners of mental health services. The structure of commissioning changed during the decade, moving from District Health Authorities, through to a combination of Health Authority and fund-holding GPs, to the development of Primary Care Groups and Primary Care Trusts. The common theme was that service commissioners wanted to ensure that they got what they paid for.

Since commissioners pay for structure (*e.g.* hospitals, mental health staff) and processes (*e.g.* Mental Health Act assessments, specific interventions), routinely collected data focussed on this 'activity data'. The result was that most if not all data collected for local (commissioner) or national return focussed on aspects such as where, when, how often and for how long patients were seen, either in individual meetings or over extended periods ("consultant episodes"). Notably absent was any consideration of outcome, either indirectly though identifying whether evidence-based care was provided, or directly through assessing whether the care provided resulted in any benefit for the patient receiving it.

In the late 1990s, a new emphasis on quality and outcome began to emerge. This was prompted by a number of influences:

- the focus on improved patient involvement in "Modernising Mental Health Services" (Department of Health, 1998a)
- the setting of public health targets for mental illness in "Our Healthier Nation" (Department of Health, 1998b)
- the careful synthesis of available evidence and setting of standards in the National Service Framework for Mental Health (Department of Health, 1999)

- the move towards clinical guidelines exemplified by the National Institute for Clinical Excellence
- the development and national implementation of a Mental Health Minimum Data Set by 2003 (Glover, 1997)
- the emphasis on reducing health inequalities (Acheson, 1998).

These initiatives culminated in a policy commitment to introduce routine use of outcomes in mental health in "Mental Health Information Strategy" (Department of Health, 2001). The primary rationale is that treatment outcomes need to be assessed if standards and targets are to be monitored and met. This injunction presents a substantial challenge to current practice. A systematic review of the use of outcome measures by consultant psychiatrists in 2000 found that outcome measures are used routinely for measuring clinical change over time by only 11% of psychiatrists in depression or anxiety, by 6.5% in psychosis, by 8.8% in cognitive impairment and by 4.7% in drugs or alcohol work (Gilbody, House & Sheldon, 2002a). This contrasts markedly with the high level of activity data found in this survey to be collected routinely, such as 86.2% of Trusts who collect data on duration of admission.

In the absence of sound empirical evidence, policy initiatives emphasising routine collection of outcome data run the risk of leading to ill-conceived and haphazard attempts at implementation, which will consume valuable resources, such as staff and patient time, for no evident benefit to the patient. There is some evidence of this happening already in North America, which is further ahead than the United Kingdom in implementing routine use of outcomes. Benjamin and colleagues reflect that the expectations of American policy-makers and service commissioners are that assessment of outcome will '*not be "too expensive", not show that the most expensive therapy is best, be easily comprehensible, address the things patients consider important, and, most importantly, save money*' (Benjamin, Perfetto & Greene, 1995, p. 305).

3.2.4 Routine use of outcomes at different levels

Overall, these reviews of current practice indicate that routine use of outcomes has been focussed on organisational-level needs, including national commissioning and local management of services. The mental health matrix proposed by Thornicroft and Tansella (1999) provides a framework for considering levels of outcome. They distinguished between the country / regional level, the local level and the patient level.

For each of these three levels, different types of outcome are possible. At the country / regional level, outcomes relate to secondary prevention (reducing relapse) and tertiary prevention (reducing the suffering consequent upon symptoms). Outcomes at this level include rates (*e.g.* suicide, homelessness, unemployment, inappropriate imprisonment) and sentinel events such as special enquiries. At the local level, outcomes can be extracted from country / regional level outcomes (*e.g.* standardised mortality ratios for the catchment area), can be measured locally, or can be collected through aggregating patient level data. The challenges of aggregating patient level data are that the information has to be routinely collected, recorded, available for aggregation, and sufficiently epidemiologically representative to be generalisable. In practice, none of these challenges are met. Outcomes at the country / regional and the local level have been the main focus in Australia, the United States and England.

The assumption underpinning the use of country / regional and local outcomes is that better commissioning, organisation and management of services will indirectly bring benefits to patients, through leading to improved quality of care. Little attention has yet been paid to patient-level outcomes, even though "*the primary purpose of mental health services is to optimise outcomes for individual patients*" (Thornicroft & Tansella, 1999, p. 96).

The FOCUS Study will investigate the use of patient-level outcome data to inform care. This approach will be termed **routine outcome assessment**, defined as the ongoing measurement and use of outcome data to inform decisions about whether to continue, change or curtail treatment. Emerging evidence suggests that routine outcome assessment could lead to benefits for individual patients. This evidence is now reviewed.

3.3 Routine outcome assessment may directly benefit patients

Two sources of preliminary evidence indicate that routine outcome assessment could be of benefit to patient care: its potential to promote individualised care in psychotherapy, and its potential to promote reflective practice in adult mental health services.

3.3.1 Routine outcome assessment in psychotherapy

There is evidence from the psychotherapy literature that routine outcome assessment results in a beneficial focus on specific outcomes in planning treatment approaches (*e.g.*

Marks, 1998; Lambert, Whipple, Smart et al, 2001). Asking clinicians to assess a particular outcome sensitises them to that outcome, highlights aspects which may not previously have been assessed, and shapes clinical behaviour towards influencing the outcome. It is plausible, though untested, that asking patients to assess particular outcomes may have similar effects in shaping their contribution to clinical care. For instance, it may prompt the patient to bring up specific and previously undisclosed problems.

Furthermore there is emerging evidence from cognitive therapy, a psychological intervention which was developed for use initially in depression (Beck, Rush, Shaw et al, 1979) but has now been found to be effective in modified form with a wide range of psychiatric disorders (Roth & Fonagy, 1997). One technique used in cognitive therapy is **patient-based outcome monitoring**, in which a target outcome is monitored by the patient at each therapy session, with the results charted over time and shared with the therapist (*e.g.* Lambert et al, 2001).

The focus in this discussion will be on using patient-based outcome monitoring to inform the care of individual patients, but it should be noted that such data can also be aggregated to inform local level decision-making. For example, patient profiling involves aggregating comparisons between the expected and the observed progress for individual patients within a service, to inform service development (Howard, Moras, Brill et al, 1996). Aggregation of routine outcome data has its critics. Bilsker and Goldner (2002) suggest that informant bias can negate the value of such data, and (as will be argued in more detail in Section 3.8.4.1) there are concerns about its representativeness (Dunn, 2001).

Patient-based outcome monitoring has been developed in part as a method for increasing collaboration and engagement, with the intention of indirectly improving outcome. There have been few attempts to evaluate patient-based outcome monitoring as a discrete intervention intended to *directly* improve outcome. One exception is the randomised controlled trial by Lambert and colleagues (2001) involving 609 university counselling service clients. The intervention group (n=307) completed the Outcome Questionnaire (Lambert, Hansen, Umphress et al, 1996) on a weekly basis, and feedback on the scores was provided to the therapist. Feedback comprised a graph showing scores over time, together with a colour coding to indicate expected progress.

The colour was based on the initial score, number of sessions and amount of change. The control group (n=302) completed the same assessment weekly, but with no feedback to the therapist. The primary hypothesis was that patients who were not on target for improvement would do better when their therapists received feedback than when their therapist did not receive feedback. This hypothesis was confirmed – nearly twice as many patients who were not on target for improvement experienced clinically significant improvement following feedback (26% versus 16%), and one third as many deteriorated (6% versus 23%). However, this study was undertaken in North America with people experiencing less severe mental health problems, which limits the generalisability to adult mental health services.

3.3.2 Routine outcome assessment in adult mental health services

The measurement and use of clinical outcomes to inform routine clinical care has moved from being a cognitive therapy technique for use in psychotherapy to being evaluated when used in adult mental health services (Biggeri et al, 1996; Priebe, McCabe, Bullenkamp et al, 2002). Possible mechanisms by which patient-based outcome monitoring could lead to improved outcome when used in adult mental health services will now be considered.

For the patient, the rationale is to set up the expectation of change, to reality test the common belief that no progress is being made, and to identify if indeed the therapy is working. Anecdotally, patients report this technique to be beneficial, both in 'feeling heard' and in contextualising any change – when the chart indicates deterioration this can be seen as a 'blip' rather than a downward spiral, and when the chart indicates improvement this reinforces the change in the desired direction. In other words, charting outcomes over time may in itself maintain and enhance improvements in those outcomes. This anecdotal evidence comes from using outcome measures in psychological therapies, but there is no reason to think these processes would not operate in adult mental health services.

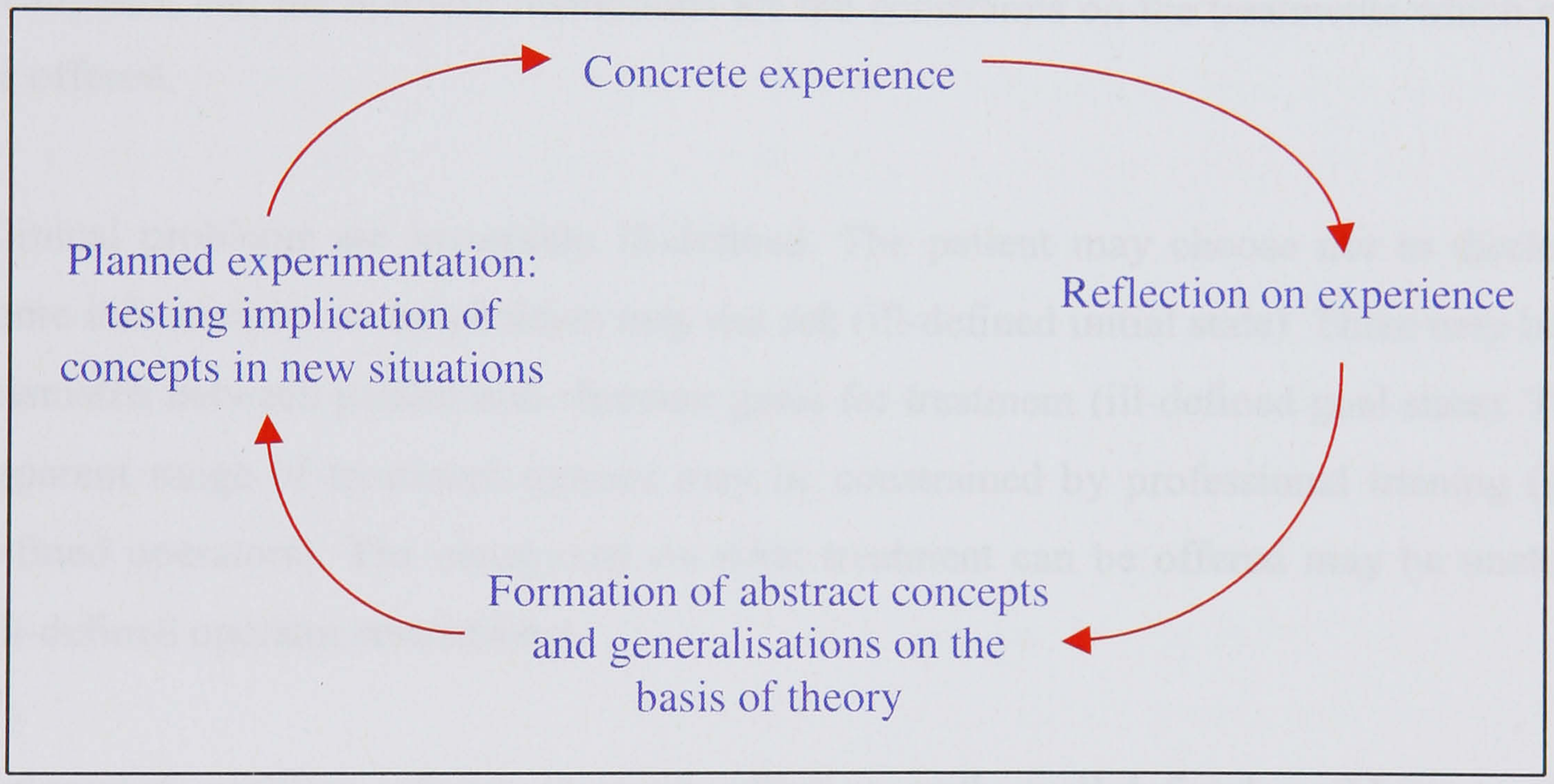
3.3.2.1 Reflective practice

The charting of patient-based outcomes can also act as a motivator for staff to re-evaluate the treatment plan where no improvement is evident. The process of developing knowledge involves both deductive reasoning (starting with an idea, developing a theory and testable hypotheses, and then gathering data to confirm or

contradict the hypotheses) and inductive reasoning (using observations to generate ideas and hypotheses, which are then tested by gathering further data) (Bowling, 1997). The feedback produced from patient-based outcome monitoring may facilitate inductive reasoning, by giving rise to questions such as "Why didn't treatment X work with patient Y? I wonder if it's because…". Such reflection has the potential to improve outcome, either through leading to changes in the content of care or by improving process issues. This approach to the planning and evaluation of treatment can be termed **reflective practice** (Kolb, 1984).

What is reflective practice? The concept can be explained with reference to Lewin's experiential learning model (Kolb, 1984), shown in Figure 3.1.

Figure 3.1: Lewin's model of experiential learning



This model of learning and action can be readily applied to clinical settings. Routine clinical practice is characterised by many demands (experiences), which require processing by clinicians (reflection), matching with theory acquired during training, and provision of further care (planned experimentation). Reflective practice involves passing around this cycle in a conscious ('reflective') manner. However, over time clinicians become skilled in these processes, and develop internal action plans based on a wide range of experiences. Using the terminology of cognitive psychology, declarative (factual and academic) knowledge becomes transformed into procedural (skills-based) knowledge through practice (Anderson, 1983). Once procedural knowledge has been developed, problem solving requires accurate pattern recognition (diagnosis or formulation) to activate the appropriate procedural knowledge, stored as a

problem-solving template. This approach can be termed **automated problem-solving**. Reflective practice and automated problem-solving lie at opposite ends of an awareness continuum, with automated problem-solving characterised by minimal conscious processing, and reflective practice by maximal conscious processing (or reflection).

3.3.2.2 Types of problem

Problems lie on a continuum from well-defined to ill-defined (Kahney, 1986). Well-defined problems are characterised by the availability of full information about the *initial state* of the problem, the *goal state*, the legal *operators* (things that can be done to solve the problem) and the *operator restrictions* (constraints on the application of operators). The less adequate the available information is, the more the problem becomes ill-defined. In the context of clinical decision-making, the initial state is the assessment, the goal state is the intended outcome of care, the operators are potential treatments, and the operator restrictions are the constraints on the treatments which can be offered.

Clinical problems are invariably ill-defined. The patient may choose not to disclose some information, or the clinician may not ask (ill-defined initial state). There may be a mismatch between patient and clinician goals for treatment (ill-defined goal state). The apparent range of treatment options may be constrained by professional training (ill-defined operators). The constraints on what treatment can be offered may be unclear (ill-defined operator restrictions).

Automated problem-solving is most effective with well-defined problems – the approach can be rapidly applied, without time-consuming thought and reflection, to many well-defined problems. It is asserted here that clinical decision-making relies mainly on this approach. For example, the culture of mental health care is to value quantity of patients seen, with caseload in generic community mental health teams often being used as a proxy measure for work effort. Sustaining a high caseload requires the frequent use of automated problem-solving, to ensure that defensible decisions are made as often as possible. There are limited incentives to reflect on experience – the nearest to this is when care plans are reviewed, but this is often in the context of a multidisciplinary review, rather than during ongoing clinical work. Thinking or reflection time can become seen as a dispensable luxury in meeting otherwise

overwhelming clinical demands. Hence automated problem-solving can become the dominant approach.

Unfortunately, automated problem-solving does not work well with ill-defined problems, which are exactly the type of problem which often arise in mental health practice. There is evidence that well-defined and ill-defined problems require different approaches to problem-solving. Schraw and colleagues showed that performance on well-defined problems was independent of performance on ill-defined problems (Schraw, Dunkle & Bendixen, 1995), and Jausovec (1987) identified differences in verbal protocols (thinking-aloud transcripts recorded during a task) for the two forms of problem.

Ill-defined problems require reflection to identify an optimal solution. This means that best quality care which maximises outcome needs more reflective and less automatic practice. If mental health services are to maximise quality of care, then there needs to be active encouragement, facilitation and rewarding of reflective practice in the work of individual clinicians. In summary, it is asserted that automated problem-solving is more time-efficient but less effective than reflective practice for the kind of problems which occur clinically. Therefore, it is worth investigating the best balance between the two problem-solving styles.

An analogous process to reflective practice can be proposed for patients. They will have views about the causes of their own problems (initial state), what would constitute a successful outcome (goal state), what help (operators) will be beneficial, and the relative importance of, for example, medication and side-effects (operator constraints). Patients differ in the extent to which they want to be involved in decisions about their care, and preferred style may change over time – becoming more involved as they become empowered to express their views (Guadagnoli & Ward, 1998). Supporting reflection and discussion by the patient will increase this involvement. It is a testable question as to whether increased involvement in decision-making is associated with improvements in process and resulting outcomes of care.

The routine collection and feedback of outcome data may help both staff and patient to think about the care being provided. This may foster reflective practice for staff and increased involvement for patients in decision-making about their treatment and care.

There have been a range of attempts to implement routine outcome assessment. In an attempt to synthesise and learn from these previous experiences, the second part of the systematic review reported in Chapter 2 is now described.

3.4 Results – principles for routine assessment of outcome

The literature was reviewed to identify principles which have been proposed for implementing routine collection of outcomes, using the method described in Section 2.2. Seven distinct frameworks for routinely assessing outcome were identified, and these are shown chronologically in Table 3.1, with individual principles grouped and ordered by degree of consensus.

Table 3.1: Proposed frameworks for routinely assessing outcomes in mental health services

Principle	Study:	1	2	3	4	5	6	7	Number of studies endorsing this principle
1. Standardised measures should be used		✓	✓		✓	✓	✓	✓	6
2. Relevance to informing practice should be emphasised		✓		✓	✓	✓		✓	5
3. Multiple perspectives should be used		✓		✓		✓	✓	✓	5
4. Standardised methods should be used		✓		✓		✓		✓	4
5. Data collection should be cheap and simple		✓			✓	✓		✓	4
6. Measures should be relevant to the patient group		✓			✓	✓		✓	4
7. Treatment received should be characterised			✓		✓	✓	✓		4
8. Feedback should be quick, easy and meaningful		✓	✓					✓	3
9. Aggregated data should be comparable with benchmarks		✓	✓					✓	3
10. Meaning of measures should be comprehensible		✓			✓			✓	3
11. Data should be collected longitudinally				✓		✓	✓		3
12. Casemix (*e.g.* diagnosis) should be assessed					✓	✓	✓		3
13. Measures should show means/processes of change		✓						✓	2
14. Measures should fit with psychopathology theories		✓						✓	2
15. Outcomes chosen should be multidimensional				✓		✓			2
16. Costs should be included				✓					1
17. Data on treatment leavers should be collected						✓			1
18. Individual utility differences should be considered				✓					1

[1] Ciarlo, Edwards, Kiresuk et al, 1981, [2] Ellwood, 1988, [3] Attkisson, Cook, Karno et al, 1992, [4] Burnam, 1996, [5] Shern & Flynn, 1996, [6] Smith, Rost, Fischer et al, 1997, [7] Stedman, Yellowlees, Mellsop et al, 1997.

The final principle of considering individual utility differences means that the outcome domains chosen should be weighted individually, since different patients may attribute varying levels of importance to particular outcome domains. Resulting data cannot then be easily aggregated for comparison between patients.

The authors of studies 1 to 6 work in North American institutions, and of Study 7 in Australia. Five studies were conducted under the auspices of national bodies – the National Institute for Mental Health (studies 1 and 3) and the National Alliance for the Mentally Ill (5), university departments (6) and Government departments (7), and two by individuals (studies 2 and 4). Studies 1 and 5 are based on the findings of task forces, studies 3 and 6 on expert panels, study 7 on a literature review (strongly based on [Green and Gracely, 1987] and study 1), and studies 2 and 4 on personal expertise.

Before considering the implications of these findings, the methodological quality of the systematic reviews reported in Chapter 2 and 3 will be assessed.

3.5 Internal validity of the systematic review

How internally valid is the review? The internal validity can be considered in terms of the extent to which the review meets criteria outlined in the QUOROM Guidelines (Moher et al, 1999). Apart from advice relating to the Abstract, the QUOROM Guidelines identify twelve components to consider in assessing quality of reporting. The quality of the current review will be considered against each QUOROM component in turn:

1. QUOROM Guidelines for the Title state "*Identify the report as a meta-analysis [or systematic review] of RCTs*". This criterion was **fully met** in the publication resulting from this review (Slade, 2002).

2. Guidelines for the Introduction state "*Describe the explicit clinical problem, biological rationale for the intervention, and rationale for review*". The rationale for routine assessment of outcome was provided in Section 3.3, and the timeliness of this review was indicated in Section 3.2.3. This review unusually covered two areas, since the literature on outcome domains and implementation of routine assessment of outcomes proved impossible to disaggregate through search strategies. This criterion was **fully met**.

3. Guidelines for the Methods (Searching) state "*Describe the information sources, in detail (e.g. databases, registers, personal files, expert informants, hand-searching), and any restrictions (years considered, publication status, language of publication)*". Table 2.2 describes in detail the electronic databases and research registers which were accessed, and all restrictions are explicitly stated in Section 2.2.2. This review is concerned with conceptual topics, so the known publication bias against negative findings (*e.g.* Sterling, Rosenbaum & Weinkam, 1995) may not be as important as in a review of clinical trials. However, three means of accessing unpublished and grey literature (conference presentations, expert informants and internet search) were described in Section 2.2.3. This criterion was **fully met**.

4. Guidelines for the Methods (Selection) state "*Describe the inclusion and exclusion criteria (defining population, intervention, principal outcomes, and study design)*". The inclusion criteria could not be formally specified beyond the conceptual level, since relevant frameworks and outcome domain proposals appeared in different contexts. The main exclusion criteria used in practice were listed in Section 2.2.4. This difficulty in constructing a precise search strategy for a non-quantitative search has been acknowledged by systematic review experts: "*When searching for qualitative research for the purpose of systematic reviews, it is often not practicable to construct strategies to capture the many ways in which such research may be described*" (Khan, ter Riet, Popay et al, 2000, p. 29). This criterion was **partly met**.

5. Guidelines for the Methods (Validity assessment) state "*Describe the criteria and process used (e.g. masked conditions, quality assessment, and their findings)*". Assessing the quality of publications was problematic, due to their diversity. There is currently no consensus regarding how to rank non-quantitative research. Some commentators suggest that no ranking is possible and each article needs to be considered on its own merits (*e.g.* Howe & Eisenhart, 1990; Buchanan, 1992). Other commentators have identified frameworks for judging quality of quantitative research (*e.g.* Seale & Silverman, 1997; Oakley, 2000), although these relate in the main to methodological standards, rather than the evaluation of conceptual work. The requirement for some form of theoretical basis for outcome domains was a minimum quality assurance approach. This criterion was **partly met**.

6. Guidelines for the Methods (Data abstraction) state "*Describe the process or processes used (e.g. completed independently, in duplicate)*". The process was described in detail in Section 2.2.4, and this criterion was **fully met**. However, the reliability would have been enhanced through either duplicate reviewing or the use of more than one reviewer.

7. Guidelines for the Methods (Study characteristics) state "*Describe the type of study design, participants' characteristics, details of intervention, outcome definitions, etc. and how clinical heterogeneity was assessed*". Characterising conceptual proposals is problematic, shown by the difficulties in formally specifying inclusion and exclusion criteria. The intended type of study was described in Section 2.2.2 – related to routine assessment of outcome in adult mental health services. In practice, included publications often were not clear about their remit, and hence were difficult to characterise. This criterion was **partly met**.

8. Guidelines for the Methods (Quantitative data synthesis) state "*Describe the principal measures of effect (e.g. relative risk), method of combining results (statistical testing and confidence intervals), handling of missing data; how statistical heterogeneity was assessed; a rationale for any a priori sensitivity and subgroup analyses; and any assessment of publication bias*". Most of these criteria do not apply to this review of two conceptual topics, and there is not yet consensus about best practice in synthesising non-quantitative research: "*There are no formal procedures available to aid narrative synthesis of findings from qualitative studies within the context of a systematic review*" (Deeks, Khan, Popay et al, 2000, p.70). However, efforts were made to tabulate the frameworks in Tables 2.3 and 3.1, and to identify emergent categories of outcome domains. An attempt was made to ensure that this criterion was **met to the extent possible**.

9. Guidelines for the Results (Trial flow) state "*Provide a meta-analysis profile summarising trial flow*". The trial flow profile is intended to indicate the overall sample, together with the points of and reasons for attrition. As shown in Table 2.2, a total of 6,357 potentially relevant publications were identified electronically by the first search. Most were excluded through initial screening of study title, but no quantitative record was kept of either further additions (through accessing the grey literature or searching reference lists), or numbers excluded at each stage (removal

of duplicates, initial screening, removal following retrieval). The reason for this omission was that it was initially thought that such information would only have value if the review was a quantitative meta-analysis. However, as the review progressed it became clear that some synthesis was possible, and that the attrition rate would in any case have been of interest for identifying the key points of exclusion. This criterion was **not met**.

10. Guidelines for the Results (Study characteristics) state "*Present descriptive data for each trial (e.g. age, sample size, intervention, dose, duration, follow-up period)*". The publications were characterised in Sections 2.3 and 3.4, including authorship, funders, and method for frameworks, and authorship, focus (*e.g.* treatment outcomes) and patient group. The lack of an accepted means of characterising conceptual publications means that this criterion was **partly met**.

11. Guidelines for the Results (Quantitative data synthesis) state "*Report agreement on the selection and validity assessment; present simple summary results…; present data needed to calculate effect sizes*". The means of identifying the seven proposed outcome domains was difficult to specify beyond describing them as 'emergent', and other reviewers might have arrived at a different set, so this criterion **was partly met**.

12. Guidelines for the Discussion state "*Summarise key findings; discuss clinical inferences based on internal and external validity; interpret the results in light of the totality of available evidence; describe potential biases in the review process (e.g. publication bias); and suggest a future research agenda*". All of these points are addressed in the Discussion sections of Chapters 2 and 3. This criterion was **fully met**.

Overall, five of the QUOROM Guidelines were fully met, five partly met, one met to the extent possible, and one not met. Some criteria were not met because of the difficulty in translating guidance intended for clinical trial reviews into guidance for a review of conceptual issues. The largest threat to internal validity was the difficulty in formally specifying the inclusion and exclusion criteria.

3.6 External validity of the systematic review

Is this systematic review externally valid? Six of the seven outcome domain proposals and eleven of the fifteen proposed sets of principles come from North American authors, reflecting that purchaser-driven pressures have stimulated more activity in routine collection of outcome data there than anywhere else. The findings of this review are therefore of most relevance to North American settings. However, the work from other countries (Australia, England and Italy) was consistent with the North American work, and no authors identified specific cultural factors which would impact on the topics of this review. The external validity is difficult to establish without a broad range of empirical data, but there is no reason to think that using the seven emergent categories of outcome domains and the identified implementation principles as a starting point for routine collection of outcome data would be unwise.

Potential biases in this review have been reviewed in discussing its internal validity, and specific concerns have been highlighted at the points of selecting studies, data extraction, and data synthesis. A summary view might be that the review identified most of the well-known conceptual work in this area, but could be methodologically improved to increase reliability of data capture and abstraction.

Future replication of this review would benefit from using multiple reviewers to allow inter-rater reliability to be investigated, and more focussed inclusion and exclusion criteria, drawing on best practice in reviewing conceptual and other non-qualitative research. The trial flow profile should be investigated, which will allow a better characterisation of reasons for exclusion.

3.7 Discussion of principles to inform the routine assessment of outcome

The review identified principles underpinning the routine assessment of outcome. It was possible to identify some consensus about the key principles, especially relating to the need to use standardised measures and consider multiple perspectives, and to ensure that outcome information informs practice. These findings will now be used to propose an evidence-based approach to implementing routine outcome assessment in adult mental health services.

A number of difficulties in routine assessment of outcome have been reported, including lack of appropriate outcome measures, lack of time, lack of incentives (financial and

professional) to offset the costs of monitoring outcome, and lack of expectations from senior staff that junior staff collect outcome data (Huxley, 1998; Marks, 1998; Slade, Thornicroft & Glover, 1999d; Walter, Cleary & Rey, 1998). Purely in terms of the assessment process, there is a lack of consensus regarding what outcome domains to include, who to ask when assessing, and what assessment measures to use (Clifford, 1998). For mental health services which intend to assess outcomes routinely, the large amount of literature, mostly from North America and orientated towards assessment information required by service commissioners, may inhibit careful consideration of important issues. The temptation may be to either do nothing, or to start by deciding which specific outcome measure to use. However, a better place to start is by making conceptual decisions about the purpose and (subsequently) the method of assessing outcomes routinely.

3.8 Four-step approach to assessing outcomes routinely

The results of this review can be synthesised to give an evidence-based method for assessing outcomes in routine adult mental health services. A four-step approach to best practice is proposed. Each step will be described in turn.

3.8.1 Proposed Step 1: Develop the model

The first step is to come to a view about what benefits are intended from the assessment of outcome. This involves the development of an explicit model of the processes involved in attaining these benefits, with each component of the model described as clearly as possible. The FOCUS Study is an example of this process, since it involves the development of a model (described in Chapter 4), which is then evaluated using the method described in Chapter 6. For routine clinical settings, the process will be less formal. Nonetheless, it will be important to specify in advance the goals of assessing outcomes, and to give early consideration to how the level of success in meeting the goals will be evaluated. This development process needs to be informed by a clear identification of what is possible within the available resources.

3.8.2 Proposed Step 2: Choose the outcome domain(s)

The second step in implementing outcome assessment routinely is to use the model developed in step 1 to identify what outcome domains are appropriate to monitor in the service. This decision may be focussed by considering the seven emergent categories of outcome domains identified in Chapter 2.

Contained within the debate on what outcome domains to consider are differing conceptual frameworks for understanding mental health problems. The issues involved are complex, and have been conceptualised in a number of ways: causation versus meaning (Bolton & Hill, 1996), idiographic versus nomothetic (Schafer, 1999), and modernist versus post-modernist (Bracken & Thomas, 2001). The framework which will be described here to illustrate the issue has been proposed by Long and Dixon (1996). They identified a spectrum of approaches to understanding mental health problems, ranging from patient-defined to professionally defined. At the extreme of the patient-defined end of the spectrum lies an understanding which emphasises the importance and uniqueness of individual experience, and accords no value to comparison of one person with another. The philosophical base for this understanding is post-modernism (Laugharne, 1999), and notions of consumerism, choice and user involvement are embraced. Correspondingly, the concepts of paternalism, professional expertise and mental illness are rejected. At the extreme of the professionally defined end of the spectrum lies an understanding which emphasises the importance of using scientific knowledge to make sense of abnormal mental experiences, and accords no value to the meaning attached by the patient to these experiences. The philosophical base for this understanding is positivism, and notions of evidence-based care, clinical judgement and compliance with treatment are embraced. Correspondingly, the concepts of patient choice about treatment and the idea of a continuum of mental health problems are rejected.

Most clinical practice, of course, takes place within these extremes, but the point on the continuum will influence the outcome domains selected for routine assessment. A mental health service operating towards the patient-defined end of the spectrum will be more interested in outcome domains related to individual phenomenological change, and how the health care service is experienced. By contrast, a mental health service operating towards the professionally defined end of the spectrum will be more interested in outcome domains related to symptoms and functioning, and in ensuring that interventions are based on research evidence.

Decisions about outcome domains will reflect the values of the stakeholders in the mental health service. For example, Andrews and colleagues (1994) exclude handicap measures because they are "confounded by cultural factors" (p. 24), yet degree of handicap is undoubtedly of central importance to patients. The decision to include

outcome domains which are susceptible to cultural bias will reflect the relative importance of professional concerns for reliable and valid measures, versus patient-defined concerns for outcome domains with high personal relevance.

Overall, it was striking in this review that there was relatively little academic literature espousing frameworks or outcome domains located at the patient-defined end of this spectrum, with the exception of the work by Campbell (1996; 1998). This may be because the majority of people who have the prerequisite academic training to articulate a theoretically-grounded framework are themselves professionals, or because people whose opinions are at the patient-defined end do not use academic publications to disseminate their views. In any event, the choice of which outcome domains to measure reflects, and therefore should be informed by conscious consideration of, the values and philosophical base of the mental health service. Considering the proposal by Campbell (1996) in Table 2.3, the language of patient-defined outcome domains differs from professionally defined outcome domains in orientation (increased concern with negative aspects of care, such as dissatisfaction and iatrogenesis), focus (less on intrapsychic variables or symptoms, more on individual perceptions and the process of service delivery) and type of jargon (*e.g.* Personhood, Empowerment, Recovery). The outcome domains chosen for consideration implicitly represent values and conceptual frameworks.

3.8.3 Proposed Step 3: Resolve the technical issues

The third step is then to consider how these outcome domains will be measured within the service. A starting point can be found in the principles identified in Table 3.1, which identify general principles which could underpin attempts to introduce and maintain routine assessment of outcome. In addition, further issues which need consideration when introducing routine outcome assessment are shown in Table 3.2, amended from Trauer (1998) and Fitzpatrick (1998).

Table 3.2: Checklist of issues to consider once the outcome domain is decided

Issue	Example question to be answered
Care versus cure	What constitutes a good outcome for a patient who is not expected to recover?
Intervention-dependent versus independent	Is the goal to show that the treatment caused improvement, or just to show that improvement occurred (without reference to treatment)?
Structures and processes versus outcomes	Whose outcome is being considered? Cost containment, adherence to clinical governance protocols, reduced symptomatology and reduced visibility of the mentally ill are all outcomes from different perspectives.
Index patient versus significant others	Is the focus just on outcome for the patient, or also for their relatives or carers (in their own right)?
Individual versus organisational	Is the focus on individual change, or aggregating data to investigate changes in groups?
Direct versus indirect measures	Are direct measures (*e.g.* from the patient or carer) or indirect measures (*e.g.* from staff or service usage) to be favoured?
Objective (measures) versus subjective (meanings)	Is equal weighting given to externally observable measures and private, self-reported experiences of the patient?
Global versus multiple item measures	Are global (single-score) or multiple item measures preferred?
Generic versus condition-specific measures	Are generic measures (applicable to broad groups) or specific measures (for highly characterised conditions) preferred?
Individualised versus standardised measures	Are individualised (tailored to the individual) or standardised measures (which can be compared to group norms) preferred?
Significant time points versus fixed intervals	Should assessment be undertaken at "important" times during the patient's pathway through care, or at pre-determined fixed time periods?

For example, consider the first issue of care versus cure. What does it mean if patients using the service do not appear to be improving, or if the service is under-performing in relation to other comparable services? This may occur for at least six reasons: casemix, epidemiology, data completeness, disorder course, intervention effectiveness, and clinician competence. Considering each in turn:

1. Services targeted at the 'severely mentally ill' may show worse outcomes than services targeted at patients with less severe mental health problems, for whom health gain is greater
2. Epidemiological characteristics and levels of psychiatric morbidity within the catchment population will vary between services
3. The relative completeness of the data will impact on the level of outcome – patients with complete data may differ from those with missing data (Young, Grusky, Jordan et al, 2000)
4. Where the natural course of the disorder involves progressive deterioration, reducing the speed of decline may (in the absence of appropriate controls) be indistinguishable from worsened outcome
5. It may simply be that, given current levels of knowledge, the variance in distal outcome domains such as quality of life which can be attributed to mental health interventions is relatively low
6. Clinician competence may impact on the outcome.

One reason that clinicians may be reticent to engage in routine assessment of outcome is a concern that they will be exposing themselves to potential criticism, with results interpreted as solely indicating clinical competence. One solution to this issue is not to measure outcomes which are outside the control of the service. However, as noted by Campbell (1998), this approach would result in the exclusion of most quality of life and psychosocial outcome domains. Early consideration of this question will highlight the need to be wary of simplistic interpretations of any changes in outcome.

3.8.4 Proposed Step 4: Choose the outcome measure(s)

The fourth and final step is then to identify the most appropriate outcome measure or measures. Many resources exist to inform this decision, including (in order of publication) Kane, Kane and Arnold (1985), McDowell and Newell (1987), Thompson (1987), Spilker (1990), Bowling (1991), Wilkin, Hallam and Doggett (1992), Bech

(1993), Andrews, Peters and Teesson (1994), Barry and Zissi (1997), Dickerson (1997), Bartlett and Coles (1998), Thornicroft and Tansella (2000) and American Psychiatric Association (2000). The psychometric properties of measures for specific outcome domains have also been compared, for example in the domains of community functioning (Dickerson, 1997), health status (Bergner & Rothman, 1987; Essink-Bot, Krabbe, Bonsel et al, 1997), quality of life (Barry & Zissi, 1997), subjective well-being (Bartlett & Coles, 1998) and service evaluation (Green & Gracely, 1987).

3.8.4.1 Most outcome measures are not designed for routine use

Although many outcome measures have been developed, most are not suitable for routine clinical use. To understand why, it will help to consider how outcome measures are developed. Various approaches have been proposed (*e.g.* Salvador-Carulla, 1996), which all have in common that conceptual work is undertaken to ensure that the outcome domain to be measured is internally valid, followed by the development and testing of an outcome measure purporting to assess the outcome domain of interest. This process establishes the **standard psychometric properties** of the outcome measure, such as reliability, validity, internal consistency and sensitivity to change.

An emphasis on the standard psychometric properties of a measure as the *sine qua non* is entirely appropriate from a research perspective – the more invalid or unreliable a measure is, the lower will be the quality of data collected using it. It is analogous to the MRC Framework for Complex Health Interventions (Campbell et al, 2000), which proceeds from a broad theoretical view to a tight and narrowly defined approach to investigating efficacy in a definitive trial, followed by a broadening out with widespread dissemination as the final step. In developing measures, broad conceptual work is followed by rigorous evaluation of psychometric properties. If adequate standard psychometric properties are established then the measure is disseminated. This 'hour-glass' approach is widely accepted as best scientific practice.

However, this approach is based on an assumption which does not necessarily lead to assessments suitable for routine clinical use. The assumption is that once the psychometric properties are rigorously established, they will be retained when the measure is applied across similar settings and patient populations. This assumption may hold where the measure is used exclusively by researchers, who are in general trained in the use of outcome measures, skilled in their use and application, motivated to produce

high response rates, able to devote their time to data collection (during the data acquisition phase of a study), and backed up by an academic infrastructure, including high-specification computers and a culture which is highly supportive of data collection. In general, none of these features are present in routine mental health services. *In extremis*, routine outcome assessment involves asking busy clinicians with little interest or training in using outcome measures to add to their work-load by filling in forms, in a culture in which form-filling is 'dead time'.

This assumption that the standard psychometric properties of the assessment will be preserved in routine use is questionable. For example, concerns have been raised in relation to HoNOS about potential over-inclusiveness of ratings, lack of specificity and non-independence of observations (Preston, 2000), about its inter-rater reliability (Brooks, 2000), and about the impact of training (Rock & Preston, 2001). The same conclusion is arrived at from a statistical perspective by Dunn (2001), who argues that psychometric properties such as reliability "*are not fixed characteristics of a particular outcome measure*" (p. 15), but rather depend on the variability of the population. He also argues that careful consideration should be given to the routine use of outcome measures, since "*much routinely collected data will be unanalysable because it [sic] is unstructured, full of holes (missing values) and laden with subjective biases*" (p. 4).

This situation is understandable. The development of measures is time-consuming, requires particular skills, and is expensive. Most measures are therefore designed by, and intended for, researchers. The resulting measures are evaluated in terms of the standard psychometric properties, rather than their suitability for routine use.

In summary, starting with measures developed by researchers for research purposes, and then trying to use them in routine clinical settings, may not be the best approach.

3.8.4.2 Developing feasible measures

Other criteria need to be considered when the goal is to develop a measure for routine use. The assessments need to satisfy criteria for **feasibility**, in addition to the standard psychometric properties. One proposed definition is that the feasibility of an assessment indicates "*the extent to which it is suitable for use on a routine, sustainable and meaningful basis in typical clinical settings, when used in a specified manner and for a specified purpose*" (Slade et al, 1999d, p.245). In other words, feasibility is a

psychometric property of the outcome measure which relies not just on the measure itself, but also on the context (where, by whom) and purpose for which it is used. For example, all other things being equal, it is likely that less severe scores on a staff-rated social disability scale will be recorded if the results are used for performance-related pay decisions (with a reduced severity leading to money for the member of staff) than if they are used for caseload monitoring (with increased severity leading to a perception that the staff member has more disabled people on their caseload). Of course, the research answer to this bias would be to have an independent person rate the measure. In routine use, however, it will be the staff member or patient who completes the measure.

A better approach may be the development and use of methodologies which lead to measures intended from the outset for routine clinical use. What are the characteristics of feasible outcome measure? It has been proposed that feasibility will be improved by developing measures which are brief, simple, relevant, acceptable, available and valuable (Slade et al, 1999d), or applicable, acceptable and practical (Andrews et al, 1994). Some of these characteristics are features of the measure, whereas some are features of the context and purpose of use. Clearly these are not characteristics which can be easily added retrospectively. Indeed, attempts to take a long research measure and shorten it for routine clinical use have been criticised methodologically (Coste, Guillemin, Pouchot et al, 1997).

3.8.4.3 Available measures for routine clinical use

Given the shortcomings of the traditional process for producing feasible assessments, what measures are available which have been developed for routine clinical use? The measures which have been developed will be identified and characterised.

A substantial review relating to routine outcome assessment was published in 1994 by Andrews and colleagues in Australia. This review covered measures to assess symptoms, functioning, quality of life, burden and satisfaction, with inclusion criteria that a manual or published article describing the measure and its psychometric properties could be obtained. The resulting 95 identified measures were then evaluated in terms of their acceptability, applicability, practicality, reliability, validity and sensitivity to change. The five outcome measures which were identified as best meeting these criteria were the Role Functioning Scale (**RFS**) (McPheeters, 1984), the

Behaviour and Symptom Identification Scale – 32 item (**BASIS-32**) (Eisen, Dill & Grob, 1994), the Health of the Nation Outcome Scales (**HoNOS**) (Wing et al, 1998), the Medical Outcomes Study Short Form (**SF-36**) (Ware & Sherbourne, 1992), and the Mental Health Inventory (**MHI**) (Veit & Ware, 1983). The Life Skills Profile (**LSP**) (Rosen, Hadzi-Pavlovic & Parker, 1989) would have met these criteria apart from the cost for use. The six measures were then field-tested in Australia (Stedman et al, 2000), and HoNOS and MHI were identified as offering the greatest potential for widespread use.

Lelliott (2000) reviewed measures focussed on what patients want from service, and proposed criteria of being relevant to what people want from mental health services, including self-report by patients, being brief and easy to understand, having established standard psychometric properties, being culturally and ethnically sensitive, and supporting a range of uses. He identified four measures which go some way to meeting these criteria: the Manchester Short Assessment of Quality of Life (**MANSA**) (Priebe, Huxley, Knight et al, 1999), the Avon Mental Health Measure (**AVON**) (unpublished), the Carers and Users Expectations of Service – Users Version (**CUES**) (Lelliott, Beevor, Hogman et al, 2001), and the Camberwell Assessment of Need (Phelan, Slade, Thornicroft et al, 1995), which has several variants, of which the Camberwell Assessment of Need Short Appraisal Schedule (**CANSAS**) is most relevant to routine outcome assessment (Slade, Loftus, Phelan et al, 1999c).

Other measures explicitly intended for routine use and which are either in widespread use or have been developed since the two previous reviews include the Dartmouth COOP Functional Assessment Charts (**COOP**) (Nelson, Landgraf, Hays et al, 1990), the Helping Alliance Scale (**HAS**) (Priebe & Gruyters, 1993) to measure therapeutic alliance, the Threshold Assessment Grid (**TAG**) (Slade, Powell, Rosen et al, 2000) to measure the severity of mental health problems, and the Functional Analysis of Care Environment (**FACE**) (Clifford, 1999) to measure mental and physical state, daily functioning and social relationships.

These 14 outcome measures are compared in Table 3.3. Given the importance of brevity, Table 3.3 is ordered by the time needed to administer each measure.

Table 3.3: Comparison of outcome measures intended for routine clinical use

Measure	Time (mins)	Domains covered 1	2	3	4	5	6	7	Psychometric properties	Use	Rater	Respondent	Training required?	Free?	Used in UK?	Notes
TAG	1-3		Y	Y	Y	Y	Y		Slade et al, 2000	Mental health	Staff	Staff	No	Yes	Yes	Developed for routine clinical use
MANSA	3-5	Y							Priebe et al, 1999	Mental health	Interviewer	Patient	No	Yes	Yes	Amended from Lancashire Quality of Life Profile (Oliver, 1991)
HAS	(i) 2-3 (ii) 2-3							Y	Priebe & Gruyters, 1993	Mental health	(i) Patient (ii) Staff	(i) Patient (ii) Staff	No	Yes	Yes	Amended from Client Assessment of Treatment (Priebe & Gruyters, 1993)
RFS	5			Y		Y			McPheeters, 1984	Mental health	Staff	Staff	Yes	Yes	No	Not widely used, but recommended in (Andrews et al, 1994)
COOP	6	Y	Y	Y	Y	Y			Nelson et al, 1990	Primary care	Patient	Patient	Minimal	Yes	No	Low sensitivity to change
SF-36	5-10.			Y	Y	Y			Ware & Sherbourne, 1992	Depression General health	Patient	Patient	Yes	Yes	Yes	Copyright restrictions apply.
CANSAS	(i) 3-5 (ii) 3-5		Y	Y	Y	Y			Slade et al, 1999c	Mental health	(i) Staff (ii) Staff	(i) Patient (ii) Staff	No	Yes	Yes	Amended from Camberwell Assessment of Need (Phelan et al, 1995)
MHI	10-15		Y	Y					Veit & Ware, 1983	Mental health General population	Patient	Patient	No	Yes	No	Copyright restrictions apply. Lacks sufficient power to detect the full range of psychiatric problems, so not appropriate for psychiatric populations

Measure	Time (mins)	Domains covered							Psychometric properties	Use	Rater	Respondent	Training required?	Free?	Used in UK?	Notes
		1	2	3	4	5	6	7								
HoNOS	15-30		Y	Y	Y	Y			Wing et al, 1998	Mental health	Staff	Staff	Yes	Yes	Yes	Developed for routine clinical use. A patient-rated version is also available, though not widely used
LSP	20-25		Y	Y	Y	Y			Rosen et al, 1989	Mental health	Staff	Staff	No	No	No	Cost is Aus$1 per form
AVON	(i) 20? (ii) 5?		Y	Y	Y	Y			None published	Mental health	(i) Patient/ Advocate (ii) Staff	(i) Patient (ii) Staff	Yes	Yes	Yes	Developed by Mind, emphasises patient's perspective
BASIS-32	20-30	Y	Y	Y		Y			Eisen et al, 1994	Mental health	Interviewer or patient	Patient	No	Yes	No	Mainly used with in-patient populations
CUES	30?					Y		Y	Lelliott et al, 2001	Mental health	Patient	Patient	No	Yes	Yes	Only limited psychometric investigation
FACE	30?		Y	Y	Y	Y	Y		Clifford, 1999	Mental health	Staff	Staff	Yes	Yes	Yes	A component in a larger set of measures, including patient-completed measures

Where time is shown with a '?' there is no published information, so time given is an estimate.

Domains covered (Y=Yes): 1=Well-being, 2=Cognition/Emotion, 3=Behaviour, 4=Physical Health, 5=Interpersonal, 6=Society, 7=Services.

Rater = person administering assessment.

Respondent = person answering the questions. Where rater and respondent are the same, the measure is self-rated.

Numbered Rater / Respondent / Time describes measures which assess more than one view.

Evidence for use in UK was publication indexed on Medline or reviewer's knowledge.

Further implementation issues will arise from the decisions about the outcome measure(s) to be used. For example, outcome measures vary in what they cost, how long they take to administer, what training is required, who can use them, who makes the rating, *etc*. Each feature of a chosen measure will have implications for how the process of routine outcome assessment is introduced and maintained.

The review of theory in Chapters 2 and 3 applied systematic reviewing methodology to identify an evidence-based framework for categorising outcome domains. Routine assessment of outcome was then placed into the broader conceptual structure of outcomes management. Drawing on findings from cognitive psychology and mental health research, it was argued that routine outcome assessment has the potential to directly or indirectly improve outcomes for individual patients. Best practice in routinely collecting outcome information was then surveyed, again using systematic reviewing methodology. This review informed the development of an implementation approach for routine collection of outcome data. The approach has four steps: develop a feasible and testable model, identify the most appropriate outcome domain(s), address any technical issues, and choose the outcome measure(s). This approach is used in Chapter 4 to develop an evidence-based model for routine outcome assessment.

Chapter 4

A model of routine outcome assessment

4.1 Outcomes research informing the model

This chapter implements the four-step approach developed in Section 3.8. A model of routine outcome assessment is described, which is intended to be feasible for use in adult mental health services. It will be evaluated in an exploratory randomised controlled trial, described in Chapter 6.

4.1.1 Target outcome domains for improvement

In this model, the primary purpose of routine outcome assessment will be to improve outcome for individual patients. An indirect benefit may be the opportunity for aggregation of data at the programme and system levels (Burns & Priebe, 1996). However, the main focus will be on the use of outcome data at the patient level, to improve mental health care for individual patients.

The model was informed by findings from outcomes research. Quality of life and unmet needs were chosen as the target outcome domains for improvement, and will be the primary end-points for the clinical trial. They were chosen for two reasons. First, they cover five of the seven categories of outcome domains identified in Table 2.3: quality of life in the Well-being category, and needs spanning Cognition/Emotion, Behaviour, Physical Health and Interpersonal categories. Interventions at the individual level will probably not significantly impact on the Society category, and outcome domains in the Services category are concerned more with the experience of care than with actual benefit to the patient. Whilst in no way neglecting the importance of patients being satisfied with the care they are receiving, it is argued here that the importance of change in the first five outcome domain categories (Well-being, Cognition/Emotion, Behaviour, Physical Health, Interpersonal) outweighs the importance of change in the experience of care.

Of course, other outcome domains (such as recovery) also span several categories. The second reason for choosing quality of life as the primary clinical outcome was because it is increasingly seen as a key outcome domain for mental health services. As will be demonstrated in Chapter 5, a candidate predictor of quality of life is level of **unmet need**, which is defined as existing where the patient experiences a current and serious

problem, irrespective of any help given (Phelan et al, 1995). As will be discussed further in Section 5.2, unmet need is a better predictor than diagnosis, symptomatology or other social or clinical variables (McCrone & Strathdee, 1994; UK700 Group, 1999; Slade, Phelan & Thornicroft, 1998). Reducing unmet need may improve quality of life. However, unmet need not only influences quality of life, but is in itself an important outcome domain. Therefore quality of life and level of unmet need were chosen as the target outcome domains for improvement.

Outcome is influenced by both the content and process of care. In this context, **content of care** is defined as the interventions which are provided for patients. The content of care will typically incorporate biological components (*e.g.* medication), psychological components (*e.g.* cognitive therapy) and social components (*e.g.* attending a day centre). Similarly, the **process of care** is defined as how the content of care is provided. This will include consideration of therapeutic alliance, degree of negotiation and level of collaboration between staff and patient.

If it is accepted that mental health interventions have any effect, then it is tautologous that changing the content of care can improve outcome. There is also evidence that the process of care mediates outcome. This evidence is strongest for therapeutic alliance and level of negotiation. Improved therapeutic alliance has been widely found to be associated with improved outcome, for example in treatment of depression (*e.g.* Krupnick, Sotsky, Simmens et al, 1996; Weiss, Gaston, Propst et al, 1997; Zuroff, Blatt, Sotsky et al, 2000) and schizophrenia (Varvin, 1991; Svensson & Hansson, 1999), and in predicting the risk of in-patient violence during hospitalisation (Beauford, McNiel & Binder, 1997). Overall, reviews find a moderate but reliable relationship between therapeutic alliance (as rated by either staff or patient) and outcome. This relationship is not influenced by other moderator variables such as outcome measure or rater, time or type of alliance assessment, type of treatment or publication status of study (Keijsers, Schaap & Hoogduin, 2000; Martin, Garske & Davis, 2000). The level of negotiation present in the therapeutic relationship is associated with improved medication concordance (Fenton, Blyler & Heinssen, 1997). Both therapeutic relationship and level of negotiation are therefore important processes impacting on outcome, acting either directly as components of interventions or indirectly as mediators for interventions.

4.1.2 Feedback

The perspectives of staff and patients differ. This can be illustrated with reference to outcomes research into both quality of life and needs. It could be argued that by definition only the patient can assess quality of life, but some studies have attempted to capture staff perspectives using 'objective' indicators. These studies consistently report no better than moderate agreement between staff and patient perspectives (*e.g.* Sainfort, Becker & Diamond, 1996; Roder-Wanner & Priebe, 1998; Doyle, Flanagan, Browne et al, 1999; Ruggeri, Bisoffi, Fontecedro et al, 2001). Needs can be assessed by both staff and patients, and again they consistently differ (*e.g.* Slade, Phelan, Thornicroft et al, 1996; Slade et al, 1998; Wiersma, Nienhuis, Giel et al, 1998; Lasalvia, Ruggeri, Mazzi et al, 2000; Hansson, Vinding, Mackeprang et al, 2001). The perspectives of staff and patients are not interchangeable.

Since staff and patient perspectives differ, it is hypothesised that giving feedback which highlights this difference will create **cognitive dissonance** – a psychological phenomenon which refers to the discomfort felt at a discrepancy between what you already know or believe, and new information or interpretation. This cognitive dissonance will foster behaviour change leading to improved negotiation and collaboration. There is some evidence to support this – a systematic review identified 21 randomised controlled trials evaluating the impact on clinicians of providing feedback regarding health status (Espallargues, Valderas & Alonso, 2000). The patient groups involved were diverse (with 7 studies using mental health populations), and the interventions involved feedback of single and mainly generic assessments to health-care professionals. The impact of mental health feedback on diagnosis or treatment (*i.e.* the process of care) was considered, and meta-analysis indicated positive changes in diagnosis (11 studies, combined OR=1.91, 95% CI 1.28-2.83), but not in treatment (8 studies, combined OR=1.15, 95% CI 0.76-1.75). However, these studies did not involve feedback to patients, did not differentiate single from repeated feedback, and were in the main investigating non-mental health populations. The review authors note the possibility that "*...communication might be improved by the provision of reports to patients as well as to physicians...Unfortunately, there are no empirical data supporting these approaches, and well-designed studies that rigorously test innovative strategies are needed.*" (Espellargues et al, 2000, p. 181).

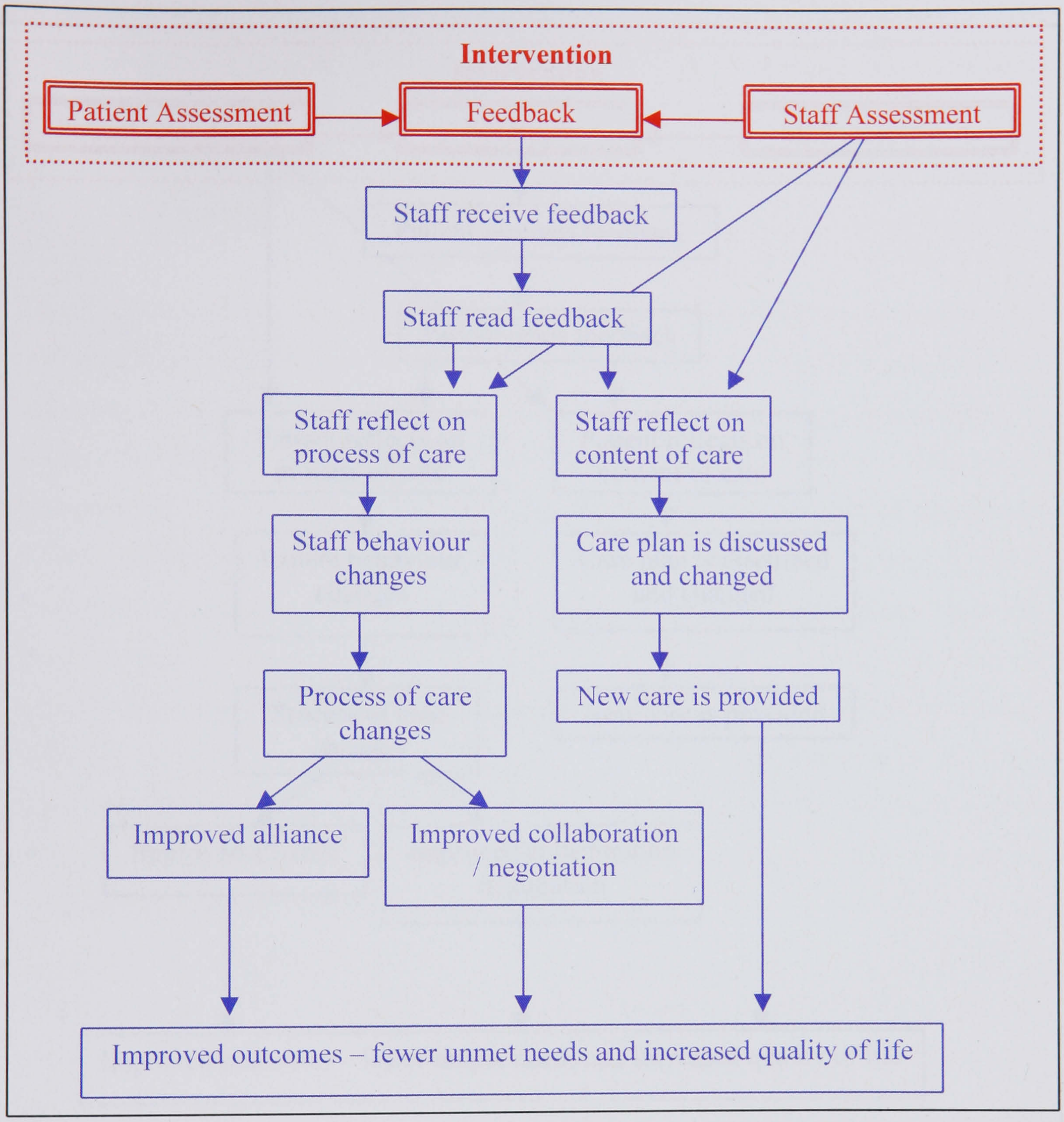
This concern is echoed in a systematic review of routine administration of health-related quality of life and needs assessment measures, which identified nine randomised or quasi-randomised studies, all in non-mental health settings (Gilbody, House & Sheldon, 2002b). The study found little evidence to support the use of these measures amongst patient populations outside of mental health services, and no randomised evidence for their use in mental health service settings. The authors conclude "*There remains an important gap in the research literature, and randomized evaluations of the effectiveness of routinely administered HRQoL [Health-related quality of life] and needs assessment tools should precede their widespread introduction*" (Gilbody et al, 2002b, pp. 1354-1355). A similar conclusion was reached in a systematic review by the same research team of the use of anxiety and depression questionnaires (Gilbody, House & Sheldon, 2001). Nine randomised studies were identified in primary care and general hospital settings, on the basis of which they concluded that routine application in these settings was not supported.

Having chosen the target outcome domains for improvement, and identified the paucity of research regarding the use of feedback in mental health services, a testable model of routine outcome assessment is now developed.

4.2 Implementation step 1: Develop the FOCUS Model

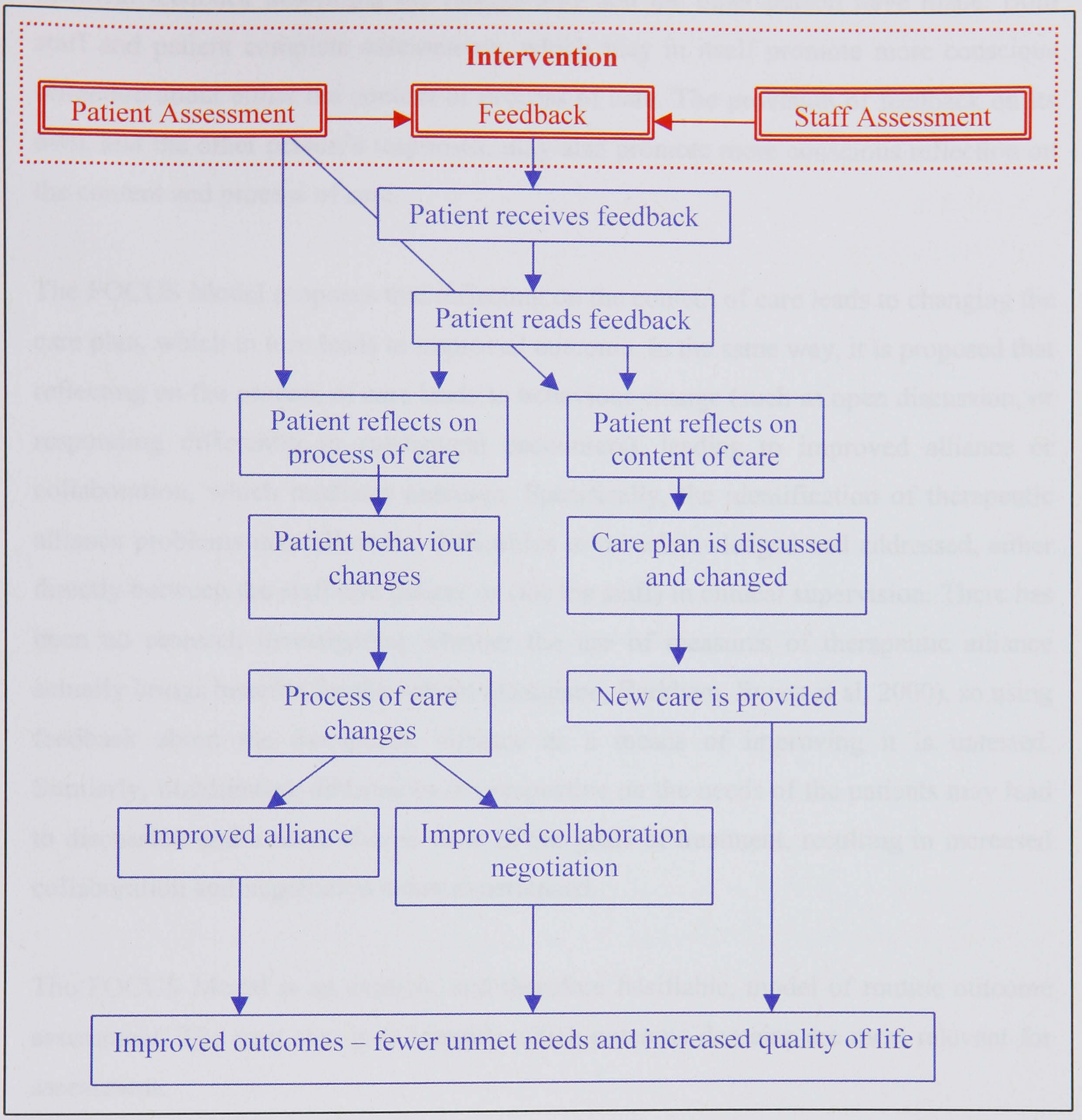
The Feedback of Outcome to Users and Staff (FOCUS) Model describes the impact of routine outcome assessment, comprising the collection and feedback of important process and outcome information. The staff version of the FOCUS Model is shown in Figure 4.1. The interventions elements are shown in red, and the intended impact of the intervention is shown in blue.

Figure 4.1: FOCUS Model of routine outcome assessment and feedback leading to improved outcome through its effect on staff



The patient version of the FOCUS Model, using the same colour scheme, is shown in Figure 4.2.

Figure 4.2: FOCUS Model of routine outcome assessment and feedback leading to improved outcome through its effect on patients



The FOCUS Model proposes that there are two pathways to improved outcome. The first pathway relates to the process of care, and the second to the content of care. Changing the content of care should directly improve outcome, and changing the process of care should indirectly improve outcome. Change in either of these pathways happens through reflection, which is facilitated by completing assessments and receiving feedback. This reflection leads to behaviour change, which in turn leads to improved outcomes.

To describe the steps in the FOCUS Model in more detail, the intervention comprises the staff and the patient being asked to complete assessments, and then each receiving identical feedback describing the ratings they and the other person have made. Both staff and patient complete assessments, which may in itself promote more conscious reflection about either the content or process of care. The provision of feedback on its own, and the other person's responses, may also promote more conscious reflection on the content and process of care.

The FOCUS Model proposes that reflecting on the content of care leads to changing the care plan, which in turn leads to improved outcome. In the same way, it is proposed that reflecting on the process of care leads to behaviour change (such as open discussion, or responding differently in subsequent encounters), leading to improved alliance or collaboration, which mediates outcome. Specifically, the identification of therapeutic alliance problems may allow the difficulties to be acknowledged and addressed, either directly between the staff and patient or (for the staff) in clinical supervision. There has been no research investigating whether the use of measures of therapeutic alliance actually brings benefits for the patient (Margison, Barkham, Evans et al, 2000), so using feedback about the therapeutic alliance as a means of improving it is untested. Similarly, highlighting differences in perspective on the needs of the patients may lead to discussion and a more shared view of the goals of treatment, resulting in increased collaboration and negotiation being experienced.

The FOCUS Model is an explicit, and therefore falsifiable, model of routine outcome assessment. The next step is to identify which outcome domains are most relevant for assessment.

4.3 Implementation step 2: Choose the outcome domain(s)

What outcome domains should be assessed routinely? The goals are to use routine outcome assessment to improve the content and the process of care. The content is changed by assessment and feedback on staff and patient assessments which relate to overall progress and benefits from treatment received. Assessment of needs, quality of life and some form of overall severity measure will promote this focus. The process is changed by assessment and feedback which produces positive changes in therapeutic alliance and in level of negotiation. Assessment of therapeutic alliance and needs (to promote discussion of what areas of life the patient has current needs in, and which are

the priority for intervention) will, in this model, beneficially impact on process. The outcome domains selected for assessment are therefore quality of life, needs, therapeutic alliance and severity of mental health problems.

4.4 Implementation step 3: Resolve the technical issues

The next step is to decide on the method of implementing routine outcome assessment. The decisions made in this step will directly inform the design of the exploratory randomised controlled trial to be described in Chapter 6. Using the principles identified in Table 3.1, the following decisions were made:

1. *Standardised measures will be used* – only standardised assessment measures will be considered for use
2. *Relevance to informing practice will be emphasised* – measures will be chosen to inform the content and process of care
3. *Multiple perspectives will be used* – staff and patient assessments will be used
4. *Standardised methods will be used* – the method of collecting data will be described, including the time interval between assessments
5. *Data collection will be cheap and simple* – outcome measures will be chosen with a strong emphasis on those which are brief and do not require formal training
6. *Measures will be relevant to the patient group* – only outcome measures which are designed specifically for mental health populations will be used
7. *Treatment received will be characterised* – measures of the treatment received will be made at baseline and follow-up
8. *Feedback will be quick, easy and meaningful* – efforts will be made to make the feedback interesting, comprehensible and relevant
9. *Aggregated data will be comparable with benchmarks* – this will not be the prime focus, since the goal of this model is benefit for the individual patient
10. *Meaning of measures will be comprehensible* – only widely-recognised outcome domains will be assessed
11. *Data will be collected longitudinally* – data will be collected on a monthly basis, and feedback will be provided on a 3-monthly basis
12. *Casemix (e.g. diagnosis) will be assessed* – the sample will be characterised by diagnosis and other standardised assessments at baseline

13. *Measures will show means / processes of change* – data will be collected relevant to each stage of the FOCUS Model, to yield information about the external validity of the model even if the intervention does not alter quality of life or unmet need
14. *Measures will fit with psychopathology theories* – measures will be chosen on the basis of the FOCUS Model, which has an explicit theoretical basis
15. *Outcomes chosen will be multidimensional* – outcome domains for the FOCUS Model span five of the seven emergent categories of outcome identified in Chapter 2
16. *Costs will be included* – the cost of implementing routine outcome assessment will be estimated
17. *Data on treatment leavers will be collected* – intention to treat analysis will be used, and full follow-up data will be collected where possible
18. *Individual utility differences will not be considered* – this principle will not be followed, since most standardised assessments do not allow utility to be assessed.

Table 3.2 identified 11 specific issues relevant to implementing routine outcome assessment. These will now be considered, and the example question given in Table 3.2 for each issue will be addressed. This will involve design decisions which will inform the FOCUS RCT.

1. Care versus cure

What constitutes a good outcome for a patient who is not expected to recover? Since some mental health patients are not expected to recover, the degree of change in outcomes between those receiving and not receiving routine outcome assessment will be compared – therefore a control group is needed. The inclusion of a measure of severity of mental health problems will ensure the focus is on reducing the impact of mental health problems.

2. Intervention-dependent versus independent

Is the goal to show that the treatment caused improvement, or just to show that improvement occurred (without reference to treatment)? The goal is to show that receiving routine outcome assessment leads to benefits for the patient, over and above those arising from treatment as usual. The actual treatment received will therefore not be assessed as part of the intervention. The primary goal of routine outcome assessment is to inform and improve future care. In the FOCUS Model it is intended that the feedback will be discussed between key-worker and patient, who will in general have a

good overview of the patient's current care. The formal assessment of previous treatment received will not have high value for informing future care.

3. Structures and processes versus outcomes

Whose outcome is being considered? The main focus will be clinical outcomes for the patient. The goal is to measure benefits for the individual patient, rather than to consider service-level issues such as cost or society-level issues such as public safety. However, therapeutic alliance and degree of negotiation are mediators of outcome, so these process measures will be included.

4. Index patient versus significant others

Is the focus just on outcome for the patient, or also for their relatives or carers (in their own right)? Although relatives and carers are important, the focus for this model is solely on the impact on the patient. Only outcome domains relevant to the patient will be considered. In the longer term, broader societal and carer outcomes will of course need to be considered.

5. Individual versus organisational

Is the focus on individual change, or aggregating data to investigate changes in groups? The primary goal is to assess the outcome for individual patients, and so data aggregation is a less important goal. Where data can be meaningfully aggregated this will be a valuable by-product, but the primary goal is to inform future care.

6. Direct versus indirect measures

Are direct measures (*e.g.* from the patient or carer) or indirect measures (*e.g.* from staff or service usage) to be favoured? A combination of assessments by the patient ('direct measures') and staff ('indirect measures') will be used, with the same assessment done by both only where it will facilitate helpful comparison and hence create cognitive dissonance. It will be important to minimise the burden on staff, since they may have to complete the outcome measure for more than one patient on their caseload.

7. Objective (measures) versus subjective (meanings)

Is equal weighting given to externally observable measures and private, self-reported experiences of the patient? Staff and patient perspectives on outcome can differ, and there is no single 'correct' perspective. Therefore both observable (staff-based) and

subjective (patient-based) measures will be used, and equal importance will be attached to both perspectives.

8. Global versus multiple item measures

Are global (single-score) or multiple item measures preferred? Global measures of mental health problem severity and multiple item measures of treatment needs will be preferred. Notwithstanding their shortcomings in terms of reduced content validity, global measures are easier to understand (i.e. have increased face validity) as outcome measures. The exception is where the assessment indicates future treatment needs, in which case item-specific measures are more appropriate.

9. Generic versus condition-specific measures

Are generic measures (applicable to broad groups) or specific measures (for highly characterised conditions) preferred? Generic measures designed for the range of people with mental health problems will be used. Generic measures make less demands on staff than condition-specific measures, since the same outcome measure is used with each patient. Minimising the burden on staff will be an important consideration in successfully implementing the intervention.

10. Individualised versus standardised measures

Are individualised (tailored to the individual) or standardised measures (which can be compared to group norms) preferred? Standardised measures are preferred over individualised measures. They are more feasible for routine clinical use (Slade et al, 1999d), and have higher face validity for staff.

11. Significant time points versus fixed intervals

Should assessment be undertaken at "important" times during the patient's pathway through care, or at pre-determined fixed time periods? Assessment will be at fixed time points. The level of practice change required for assessment at fixed time intervals is less than that required for assessment at key times during the patient's pathway through care.

4.5 Implementation step 4: Choose the outcome measure(s)

A wide range of candidate outcome measures have been developed and evaluated, as described in Section 3.8.4. The exploratory randomised controlled trial to test the

FOCUS Model will take place in routine adult mental health services, which currently have neither a culture of outcome assessment, nor any infrastructure to manage, analyse or use the resulting data. The regular rating and use of outcome data will therefore require changes in staff and patient behaviour, so it is important that the outcome measures chosen are minimally burdensome. In deciding what outcome domains to measure, the balance is between minimising the burden and maximising the value of the data collection. To maximise the likelihood of clinicians actually providing the data, this balance will need to be made at a different point to that which would be made in an efficacy trial. The following six criteria were identified to inform the choice of outcome measures to use in the intervention:

1. Based on the FOCUS Model, the measure either assesses a desired outcome (needs, quality of life) or process measure (therapeutic alliance), allows explicit comparison between staff and patient views, or is a severity measure
2. The measure has peer-reviewed published evidence of acceptable psychometric properties
3. The measure is designed specifically for a mental health population
4. The measure is brief to administer (arbitrarily chosen as an administration time of less than 15 minutes)
5. There is no charge to use the measure
6. There is no training required to use the measure.

The assessment of the 14 outcome measures identified in Table 3.3 as potentially appropriate for routine clinical use against these six criteria is shown in Table 4.1.

Table 4.1: Feasibility of selected outcome measures

Outcome measure	Specific outcome / process?	Psychometrics established?	Designed for mental health?	Brief?	Free?	No training required?	Number of criteria met
TAG	Yes	Yes	Yes	Yes	Yes	Yes	**6**
MANSA	Yes	Yes	Yes	Yes	Yes	Yes	**6**
HAS	Yes	Yes	Yes	Yes	Yes	Yes	**6**
RFS	Yes	Yes	Yes	Yes	Yes	**No**	**5**
COOP	Yes	Yes	**No**	Yes	Yes	Yes	**5**
SF-36	Yes	Yes	**No**	Yes	Yes	**No**	**4**
CANSAS	Yes	Yes	Yes	Yes	Yes	Yes	**6**
MHI	Yes	Yes	Yes	Yes	Yes	**No**	**5**
HoNOS	Yes	Yes	Yes	Yes	Yes	**No**	**5**
LSP	Yes	Yes	Yes	**No**	**No**	Yes	**4**
AVON	Yes	**No**	Yes	**No**	Yes	**No**	**3**
BASIS-32	Yes	Yes	Yes	**No**	Yes	Yes	**5**
CUES	Yes	Yes	Yes	**No**	Yes	Yes	**5**
FACE	Yes	Yes	Yes	**No**	Yes	**No**	**4**

TAG, MANSA, HAS, and CANSAS met all six criteria. These four measures were chosen for inclusion, and are described more fully in Section 6.10.1. The staff-completed measures will be TAG, staff HAS (HAS-S) and staff CANSAS (CANSAS-S). According to guidance notes, total time for completion was estimated to be 6-11 minutes. The patient-completed measures will be MANSA, patient HAS (HAS-P) and patient CANSAS (CANSAS-P), and completion was estimated to take 8-13 minutes. The intention is that comparison of the HAS assessments will focus staff and patient on the process of care, comparison of CANSAS assessments will increase collaboration and negotiation, and feedback of the MANSA, CANSAS and TAG assessments will lead to an increased focus on desirable outcomes.

This chapter has used the theory presented in Chapters 2 and 3 to develop a testable model of routine outcome assessment, and to make specific decisions about how the model can be implemented in routine clinical practice. These decisions will inform the design of the exploratory randomised controlled trial which will be used to evaluate the FOCUS Model, and whose method is described in Chapter 6.

To maximise value from the proposed randomised controlled trial, the scientific question of the relationship between the two primary outcomes for the FOCUS Model will also be considered. The intention is to develop a hypothesis which can then be tested as part of the same randomised controlled trial. Chapter 5 will investigate the relationships found in research studies between the two primary outcomes on which routine outcome assessment is intended to have an impact – unmet needs and quality of life.

Chapter 5

The relationship between patient-rated unmet need and quality of life

5.1 From model to theory

The modelling phase of the MRC Framework '*involves delineating an intervention's components and how they inter-relate and how active components of a complex package may relate to either surrogate or final outcomes*' (Medical Research Council, 2000, p. 4). This chapter addresses the question of whether the two primary outcomes for the proposed FOCUS Model – unmet need and quality of life – are independent.

A non-systematic review about the predictors of these two outcomes is reported. The aim is to identify whether there is evidence of a causal relationship between them. This leads to a preliminary conclusion that patient-rated unmet need may causally influence quality of life. This preliminary hypothesis is investigated through a retrospective re-analysis of existing data, previously collected in a routine clinical service. The re-analysis provides further evidence for a causal relationship, which informs the goals and hypotheses of the FOCUS RCT described in Chapter 6.

Existing research relating to unmet need and quality of life, will now be reviewed. The aim of this review is to establish whether previous studies are compatible with the hypothesis that high unmet need *causes* low quality of life.

5.2 Research investigating unmet need

Two meanings of the term 'unmet need' can be distinguished. The first meaning, derived from epidemiology, refers to the proportion of people who meet criteria for a disorder and do not see a clinician (Andrews & Henderson, 2000). This meaning concerns population-level needs, and is not directly relevant to this study.

The second meaning of unmet need relates to individual-level needs assessment. Related research provides data which are intended to inform the provision of care to individual patients. Two definitions of unmet need at the individual level can be identified. The MRC Needs for Care Assessment Schedule (**NFCAS**) is based on a definition of an unmet need for treatment as occurring when an effective and acceptable intervention is not delivered (Brewin, Wing, Mangen et al, 1987). It also allows two variants: an unmet need for assessment, and an unmeetable need (when a potentially

effective intervention is not available or is being refused by the patient). Within this framework, need is a normative concept to be defined by the health professional. Since its development, the NFCAS has been used in several research studies, so data on unmet need from these studies will be considered for review. However, more recent conceptualisations have challenged the notion that need should be solely professionally-defined, suggesting instead that need is a subjective concept, about which (at least) staff and patients will have valid perceptions (Slade, 1994). As defined in Section 4.1, an **unmet need** exists where the person experiences a current and serious problem, irrespective of any help given (Phelan et al, 1995). To be consistent with the CAN, therefore, the three NFCAS categories of unmeetable needs, unmet needs for assessment and unmet need for treatment are reduced to the single category of unmet need.

The literature on unmet need was reviewed by searching Medline for all articles published between 1993 and June 2001 matching the search term '(unmet need? or camberwell assessment of need).tw'. This search returned 352 matches. Reviewing the titles for relevance to mental health reduced this list to 50 matches. Reviewing the abstracts to identify those publications reporting data on the relationship between unmet need and other variables for adults reduced the list to 19. Studies excluded following abstract review were investigating psychometric properties (n=12), relevant to the epidemiological definition of unmet need (n=9), investigating factors reducing access (n=4), treatment reviews (medication for depression, "needs feedback") (n=2), or relating to other client groups – child (2), alcohol (n=1) and myocardial infarction (n=1). Reading the 19 papers reduced the list to 16, with 2 studies excluded because they were relevant to the epidemiological definition of unmet need and 1 because it reported no data on unmet needs.

The findings from the 16 identified studies, ordered by their sample size, are shown in Table 5.1. Four aspects of the design are summarised: cross-sectional versus longitudinal (studies presenting cross-sectional data from a longitudinal study were classified as cross-sectional); single-site versus multi-site, cohort (either epidemiologically representative or consecutive) versus convenience sample, and case-controlled versus uncontrolled. Unless otherwise specified, studies were single-site, uncontrolled and used a convenience sample.

Table 5.1: Studies investigating the association between unmet need and other variables

	n	Inclusion criteria	Design	Measure	Associations found between unmet need and other variables
1	708	Psychosis	cross-sectional cohort	CAN	Higher staff-rated unmet need was associated with lower quality of life, more so for African-Caribbeans. Strongest association was with basic, social and functioning needs.
2	337	Psychiatric patients	cross-sectional multi-site cohort	CAN	Level of staff-rated unmet need correlated with total HoNOS score, with correlation at least 0.5 ($p<0.001$) for domains of drugs, psychotic symptoms, company and accommodation
3	300	Schizophrenia	cross-sectional multi-site	CAN	Higher patient-rated unmet need was associated with worse quality of life and social network. Higher staff-rated unmet need associated with worse psychopathology and social network.
4	247	Psychiatric patients	cross-sectional cohort	CAN	Higher staff-rated (but not patient-rated) unmet need was associated with having a diagnosis of psychosis or personality disorder and not neurosis or other diagnosis.
5	177	In-patient	12-month longitudinal case-controlled	CAN	Admission to residential units attached to community mental health centres (community beds) resulted over time in lower staff-rated (but not patient-rated) unmet needs, compared with admission to acute units
6	133	Functional psychosis	cross-sectional cohort	CAN	Lower quality of life was associated with higher unmet need, however assessed. The relationship was stronger for patient ratings. Patient unmet need ratings were more reliable than staff ratings.
7	131	Functional psychosis	2-year longitudinal case-controlled	CAN	Higher satisfaction was associated with fewer patient-rated unmet needs at both baseline and follow-up. Baseline unmet need predicts follow-up unmet need.
8	120	Schizophrenia	cross-sectional cohort	CAN	Higher patient-rated unmet need was associated with lower quality of life, especially for company, psychological distress, daytime activities and sexual expression.
9	82	23 never, 16 current and 43 ever had mental illness	12-month longitudinal	NFCAS	At baseline, higher staff-rated unmet need associated with ever having a mental illness, physical or sexual abuse during childhood, social network problems, no intimate relationship ever, and unemployment. No conclusions about baseline predictors of unmet need at follow-up.
10	78	Psychiatric patients	cross-sectional	CAN	Level of staff-rated unmet need was highly correlated with total HoNOS score, but level of patient-rated unmet need was not correlated.
11	73	Long-term mental illness	cross-sectional	Unpublished questionnaire	Patient-rated unmet needs for information and help associated with poor quality of life. Safety, money and employment needs associated with rehospitalisation and emergency room visits
12	50	Psychosis	2-year longitudinal cohort	NFCAS and CAN	Higher staff-rated needs were associated with being single and in-patient. Unmet need was associated with a reduced follow-up rate.

	n	Inclusion criteria	Design	measure	Associations found between unmet need and other variables
13	49	Severe / enduring mental illness	cross-sectional cohort	CAN	No relationship between patient-rated unmet needs and whether receiving primary or secondary mental health care
14	49	Psychiatric patients	cross-sectional	CAN	Assessment of staff and patients differed, especially in domains of need for which there is no easily-defined service response
15	47	Functional psychosis	cross-sectional cohort	NFCAS	No relationship between staff-rated unmet need and diagnosis.
16	40	Functional psychosis	cross-sectional cohort	CAN	Drug or alcohol misuse plus functional psychosis associated with higher levels of staff-rated unmet need than psychosis-only group

Measures: NFCAS=Needs for Care Assessment Schedule (Brewin et al, 1987), CAN = Camberwell Assessment of Need (Phelan et al, 1995)

[1] UK700, 1999, [2] Slade, Beck, Bindman et al, 1999a, [3] Hansson et al, 2001, [4] Lasalvia et al, 2000, [5] Boardman, Hodgson, Lewis et al, 1999, [6] Slade, Leese, Taylor et al, 1999b, [7] Leese, Johnson, Slade et al, 1998, [8] Bengtsson-Tops & Hansson, 1999, [9] Lefebvre, Cyr, Lesage et al, 2000, [10] Issakidis & Teesson, 1999, [11] Perese, 1997, [12] Wiersma et al, 1998, [13] Barr, 2000, [14] Slade et al, 1996, [15] Wiersma, Nienhuis, Giel et al, 1996, [16] Wright, Gournay, Glorney et al, 2000.

The only studies using more than one time-point were studies 5 (Boardman et al, 1999), 7 (Leese et al, 1998), 9 (Lefebvre et al, 2000) and 12 (Wiersma et al, 1998). Studies 5 and 7 investigated approaches to in-patient and community care respectively, study 9 validated the French version of the NFCAS, and study 12 was primarily investigating the stability of needs between 15 and 17 years after the first onset of psychosis. None of these studies was therefore designed specifically to identify predictors of unmet need.

Since current evidence on the relationship between unmet need and other variables is mainly cross-sectional, only association (not causation) can be inferred. However, this review allows some tentative conclusions. There is no consistent evidence of higher levels of unmet need being associated with any sociodemographic variable, for example age, sex or ethnicity. Staff and patient ratings of unmet need are associated with different outcomes. Higher staff-rated unmet need is associated with higher social disability, lower quality of life, and a poorer social network. Higher patient-rated unmet need is associated with lower quality of life and a lower satisfaction with care. No association between diagnosis and patient-rated unmet need is found, but there is some evidence that psychosis or personality disorder is associated with higher staff-rated unmet needs than other diagnoses, and that comorbid substance abuse and psychosis is associated with more staff-rated unmet needs than psychosis alone.

5.3 Research investigating quality of life

The literature on quality of life was reviewed by searching Medline for all articles published between 1993 and June 2001 matching the search term 'Quality of life/ and Mental disorders/'. This search returned 277 matches. Reviewing the titles for those indicating data on quality of life for general adult mental health patients reduced this list to 59 matches. Reviewing the abstracts to identify those publications reporting data on the relationship between quality of life and other variables for adults reduced the list to 20. Studies excluded following abstract review were testing the impact of different services on quality of life (n=14), evaluating psychometric properties (n=9), investigating quality of life as a predictor (n=6), ethnographic (n=1) or conceptual (n=5) investigations, comparing quality of life in people with and without mental health problems (n=3), or focussing on health status (n=1). Reading the papers reduced the list to 18, with 2 studies excluded because they inferred quality of life from other variables, rather than assessing it directly. The findings from the 18 studies, ordered by their sample size, are shown in Table 5.2.

Table 5.2: Studies investigating the association between quality of life and other variables

Study	n	Inclusion criteria	Design	Measure	Associations found between quality of life (QOL) and other variables
1	4257	Homeless mentally ill	12-month longitudinal cohort	QOLI	QOL improved by 30% over time. Baseline QOL was negatively associated with depression, psychotic symptoms, intoxication, drug use and education, and positively associated with social support, ethnicity, use of services and income. Changes in these variables were associated with increased follow-up QOL.
2	1000	Primary care	cross-sectional	SF-20	More variance in QOL was accounted for by common mental disorders (especially depression, anxiety, somatoform and eating disorders) than by physical disorders.
3	620	Psychotic patients	cross-sectional multi-site	LQL	27% of variance in QOL was accounted for by a wide range of social, clinical and needs data. Separate analysis indicated that unmet needs (especially basic, interpersonal and social functioning) accounted for 20% of variance, clinical variables (especially depression and positive symptoms) for 19%, and social variables for 7%. Similar but stronger effect found for Afro-Caribbean patients.
4	592	Psychiatric patients	cross-sectional	QOLIMH	QOL and service satisfaction correlated for schizophrenia, affective disorders and adjustment disorders, but not for anxiety disorders.
5	250	Psychiatric patients	12-month longitudinal	Q-LES-Q	QOL improved over time, and out-patients had higher overall QOL than in-patients. Low baseline vulnerability score (marital status, income, health, life conditions, occupation, childhood risk factors) and in-patient status were associated with better 'social relations'.
6	252	Psychiatric patients	cross-sectional	QOLI	QOL was lower in homeless than domiciled patients, especially in living situation, finances, daily activities and family relations.
7	243	Psychiatric patients	cross-sectional	Unpublished questionnaire	58% of variance in QOL was accounted for by depression (48%), poor social and financial circumstances, and poor self-rated health.
8	237	Psychiatric in-patients	cross-sectional	?	QOL was positively associated with level of need.
9	210	Psychiatric in-patients	cross-sectional case-controlled	Q-LES-Q	Psychiatric patients had lower QOL than controls. 45% of poor QOL in schizophrenia was accounted for by depression, somatization, anergia, activation, drug-induced abnormal voluntary movements, and emotion-focused coping, and 48% of good QOL by task- and avoidance-oriented coping, perceived social support from family and significant others. Psychosocial factors have a higher impact than symptoms.
10	194	Psychiatric patients	6-month longitudinal cohort	LQL	No baseline predictors of follow-up QOL were found. Follow-up satisfaction with care and QOL were associated.
11	184	Psychiatric patients	cross-sectional	QOLI	Lower QOL was associated with "laissez-faire" leadership of services, and higher QOL with "transformational" and "transactional" styles.

Study	n	Inclusion criteria	Design	Measure	Associations found between quality of life (QOL) and other variables
12	183	Psychiatric patients	30-month longitudinal cohort	LQL	Subjective ratings constitute a separate factor from objective ratings. 55% of variance in baseline subjective QOL was explained by older age, higher service satisfaction, positive affect and greater self-esteem. 44% of variance in follow-up subjective QOL was explained by baseline subjective QOL and unemployment
13	173	Homeless mentally ill	3-month longitudinal	LSM	Higher QOL was found in people who went on to received welfare benefits, compared with those who were denied benefits.
14	165	Psychiatric patients	cross-sectional	SLDS	QOL did not differ between sexes, but increased with older age.
15	92	Psychiatric patients	cross-sectional	QOLI	QOL was higher in employed than unemployed. 20% of variance in QOL was accounted for by valuing of work, age and number of hospitalisations.
16	74	Psychiatric patients	cross-sectional cohort	?	Higher QOL was found in those living outside institutions, with 63% of variance accounted for by loneliness, satisfaction with neighbourhood, and leisure activities.
17	70	Psychiatric patients	18-month longitudinal	LQL	QOL did not differ between sexes or ethnic groups, but increased with older age. Cross-sectional associations with depression, age, life experiences and QOL domains accounted for 59% of variance. Change in depression and QOL domains accounted for 36% of the variance in QOL change.
18	49	Psychiatric in-patients	cross-sectional	QOLI	55% of the variance in QOL was accounted for by depression, social network size, verbal intelligence and social adjustment.

QOL = quality of life

Measures: QOLI = Quality of Life Interview (Lehman, 1983), SF-20 = Short-Form General Health Survey (Stewart, Hays & Ware, 1988), LQL = Lancashire Quality of Life Profile (Oliver, 1991), QOLIMH = Quality of Life Index for Mental Health (Becker, Diamond & Sainfort, 1993), Q-LES-Q = Quality of Life Enjoyment and Satisfaction Questionnaire (Endicott, Nee, Harrison et al, 1993), ? = only abstract available in English, which does not report assessment used, LSM = Lehman's Summary Measure (Lehman, 1988), SLDS = Satisfaction with Life Domains Scale (Baker & Intagliata, 1982)

Studies: [1] Lam & Rosenheck, 2000, [2] Spitzer, Kroenke, Linzer et al, 1995, [3] UK700, 1999, [4] Rohland, Langbehn & Rohrer, 2000, [5] Henkel, Schmitz, Berghofer et al, 2000, [6] Lehman, Kernan, DeForge et al, 1995, [7] Koivumaa-Honkanen, Viinamaki, Honkanen et al, 1996, [8] Hoffman, Priebe, Isermann et al, 1997, [9] Ritsner, Modai, Endicott et al, 2000, [10] Ruggeri, Biggeri, Rucci et al, 1998, [11] Corrigan, Lickey, Campion et al, 2000, [12] Ruggeri et al, 2001, [13] Rosenheck, Dausey, Frisman et al, 2000, [14] Mercier, Peladeau & Tempier, 1998, [15] Van Dongen, 1996, [16] Borge, Martinsen, Ruud et al, 2000, [17] Holloway & Carson, 1999, [18] Corrigan & Buican, 1995.

This review allows some tentative conclusions about the relationship of quality of life to sociodemographic, clinical and social characteristics. Quality of life differs between ethnic groups, though not in a consistent manner, and improves with age. This age effect will contribute to the common finding that quality of life typically improves between baseline and follow-up, regardless of intervention. No association with gender was evident in the reviewed studies, although epidemiological evidence suggests an interaction between gender and marital status for mental health – for example, recent evidence from the British Household Panel Survey (Willitts, Benzeval & Stansfield, 2004) indicates that cohabitation provides more mental health protection for men and marriage more mental health protection for women. There is good evidence that mental health problems (especially depression) reduce quality of life, more so than common medical problems. No consistent relationship with specific symptoms is found, but high satisfaction with care and low levels of need (especially unmet need) are associated with better quality of life, There is some evidence that psychological factors such as coping strategies and self-esteem impact on quality of life. Good social networks, family support, having money, and being employed and domiciled are all associated with higher quality of life. No more than two thirds of the variance in quality of life is likely to be accounted for by a model considering all these variable.

Taken together, the findings from these two reviews were not conclusive. Most identified studies used cross-sectional designs. Furthermore, a non-systematic review methodology was employed. This means that research published before 1993, studies not matching the search criteria, and studies not indexed on Medline would all have been omitted from consideration. A systematic review was not appropriate because the intention was to establish whether the available evidence was broadly compatible with the causal hypothesis, rather than to review all research findings.

The findings from these two reviews were consistent with the hypothesis that high patient-rated unmet need *causes* lower quality of life. To consider whether further preliminary evidence exists to support this hypothesis, an existing longitudinal outcomes database was retrospectively re-analysed. This analysis is now reported.

5.4 Hypotheses for confirmatory re-analysis

Bollen (1989) proposed three criteria for establishing a causal relationship: association (the putative cause and effect have temporal and spatial contiguity), direction (cause

precedes effect) and isolation (the effects of a cause are isolated from competing causes). To investigate whether high patient-rated unmet needs cause low quality of life, association involves testing whether people with high patient-rated unmet needs also have low quality of life. Direction involves testing whether quality of life improves after a patient-rated unmet need is met. Isolation involves testing whether quality of life improves when unmet needs are met, but not when (say) social networks increase. Based on the studies described in Tables 5.1 and 5.2, there is currently cross-sectional evidence for association and isolation in the relationship between patient-rated unmet need and quality of life, but no longitudinal evidence addressing whether meeting patient-rated unmet needs *causes* subjective quality of life to improve.

Two hypotheses were investigated using routinely-collected outcome data:

(i) The number of patient-rated unmet needs is cross-sectionally inversely associated with subjective quality of life for all patients, whether or not clinical and social variables are controlled for (cross-sectional isolation and association respectively)
(ii) The number of patient-rated unmet needs at baseline predicts level of subjective quality of life at one-year follow-up (longitudinal association).

5.5 Method of confirmatory re-analysis

Data presented here are part of the larger South-Verona Outcome Project (SVOP), a naturalistic longitudinal study which assesses the outcome of care provided by the South-Verona Community Mental Health Service using standardised instruments completed within routine clinical practice (Ruggeri et al, 2001). The study involves assessments of all patient in contact with the Service (as recorded in the psychiatric case register), with repeated follow-up assessments of regular attenders.

5.5.1 Confirmatory re-analysis: Participants

The study took place in South Verona (population 75,000), a predominantly urban area in north-east Italy. The South Verona Community Mental Health Service provides comprehensive integrated services, emphasising continuity of care by employing staff (excluding nurses) who work across hospital and community facilities (Ruggeri et al, 2001). The analysis reported here is based on the attenders assessed during the study period of October to December 1996, and the sub-group who were still in contact one year later (October to December 1997).

5.5.2 Confirmatory re-analysis: Procedures

Assessments took place the first or second time the patient saw a psychiatrist or psychologist (other than Casualty or liaison psychiatry contacts) during the study period. All staff were trained in the use of standardised assessments. Researchers assisted patients in completing the assessments where necessary.

5.5.3 Confirmatory re-analysis: Measures

Mental health staff completed four assessments at baseline and follow-up. The Global Assessment of Functioning (**GAF**) scale is a single-item measure of functioning from 0 (extremely severe dysfunction) to 90 (extremely good function) (Endicott & Spitzer, 1976). The Brief Psychiatric Rating Scale (**BPRS**) expanded version is a 24-item measure of symptomatology, covering anxiety/depression, positive symptoms, negative symptoms, mania and cognitive impairment (Overall & Gorham, 1988). Each item is rated from 1 (no symptom) to 7 (extremely severe symptom). The Disability Assessment Schedule (**DAS**) is an 8-item assessment of social role functioning, with each item rated from 0 (no dysfunction) to 5 (maximum dysfunction) (World Health Organization, 1998). The Camberwell Assessment of Need (**CAN**) assesses the presence of a met or unmet need in 22 health and social domains (Phelan et al, 1995; Slade et al, 1999c).

Patients also completed the CAN at baseline only. In addition, patients completed at baseline and follow-up the Verona Service Satisfaction Scale (**VSSS**) which assess 54 aspects of care from 1 (terrible) to 5 (excellent) (Ruggeri & Dall'Agnola, 1993), and the Lancashire Quality of Life Profile (**LQL**) which assesses quality of life (Oliver, Huxley, Priebe et al, 1997). Only LQL subjective measures are reported in this study, which use a 7-point Likert scale from 1 ("My life couldn't be worse") to 7 ("My life couldn't be better") for general well-being and eight more specific domains – leisure/participation, religion, finances, living situation, legal and safety, family and social relations, health and self-concept. The mean score across all nine domains was used as the LQL score. VSSS and LQL assessments relate to the previous year, and all other assessments to the previous month.

5.5.4 Confirmatory re-analysis: Methods of Analysis

Non-responder differences were tested using independent-sample t-tests. Hypothesis (i) was tested using linear regression with variables entered in blocks comprising

sociodemographic data (sex, age, and dichotomous variables for being married and being employed as opposed to unemployed, home maker, student or retired), baseline diagnosis (psychosis or not), and baseline dependent variables (BPRS, GAF, DAS, and either staff CAN-unmet and CAN-met, patient CAN-unmet and CAN-met, or both). Hypothesis (ii) was tested using linear regression analysis on baseline values (with and without baseline LQL included as an independent variable), with follow-up LQL score as the dependent variable. For both analyses, percentages of variance are adjusted R^2 statistics, B is the regression coefficient, and Beta is the coefficient when the dependent and independent variables are standardised to have unit standard deviation.

Hypothesis (ii) was also investigated using graphical modelling (Edwards, 2000; Biggeri et al, 2001). The partial correlation matrix generated by multiple regression analysis indicates the associations between all variables, with zero partial correlation coefficient indicating condition independence. Graphical modelling, by contrast, provides a simpler, visual structure of the relationship between variables, by setting non-significant partial correlations to zero. The result is a graph with nodes denoting variables, edges indicating an association between variables, and the absence of an edge indicating conditional independence (although the two variables may be marginally independent if they are connected indirectly through other nodes).

A graphical chain model of the relationship between baseline and follow-up scores for CAN and LQL was constructed, to illustrate the strongest relationships among all the variables taken together. Baseline variables were fixed within the model. The stepwise backward procedure was used to select a model, removing partial correlations not significant at p=0.01. Graphical modelling relies on the assumptions that non-linear relationships are negligible and that the relationship between variable pairs is not modified by a third variable. "Leave-one-out" residuals from the final fitted model were examined for evidence of non-normality and non-linearity, to test these assumptions (Edwards, 2000). Regression analyses were performed using SPSS for Windows 8.0.1, and estimation and fitting of the graphical model using MIM.

5.6 Results of confirmatory re-analysis

The clinical and social characteristics of the 265 patients assessed at baseline and of the sub-group of 121 long-term patients assessed at one-year follow-up are shown in Table 5.3.

Table 5.3: Clinical and social characteristics of all patients (n=265) and long-term patient sub-group (n=121) at baseline

	All patients	**Long-term patient sub-group**
Age	45.7 (s.d. 15.5)	45.8 (s.d. 15.8)
Male	95 (36%)	43 (36%)
Marital status		
Unmarried	106 (40%)	55 (46%)
Married	107 (40%)	43 (36%)
Widowed / separated / divorced	52 (20%)	23 (18%)
Living situation		
Alone	38 (14%)	18 (15%)
With family or relatives	215 (81%)	96 (79%)
Hospital / hostel	12 (5%)	7 (6%)
Employment		
Employed	97 (37%)	34 (28%)
Unemployed	36 (14%)	24 (20%)
Home-maker / retired / student	132 (50%)	63 (52%)
Diagnosis		
Schizophrenia and other functional psychosis	75 (28%)	49 (41%)
Affective psychosis	19 (7%)	9 (7%)
Depressive neurosis	94 (36%)	38 (31%)
Other neurosis	37 (14%)	11 (9%)
Personality disorder	20 (8%)	10 (8%)
Other or no psychiatric diagnosis	20 (8%)	4 (3%)
Outcome measure (range)	*Mean score (s.d.)*	
GAF (0-90)	59.0 (15.5)	56.0 (15.6)
DAS (0-5)	0.63 (0.92)	0.80 (1.01)
BPRS (1-7)	1.51 (0.47	1.57 (0.53)
staff CAN-unmet (0-22)	0.88 (1.54)	1.12 (1.60)
staff CAN-met (0-22)	2.37 (2.29)	2.68 (2.52)
patient CAN-unmet (0-22)	1.22 (2.03)	1.50 (2.36)
patient CAN-met (0-22)	1.85 (2.03)	2.10 (2.12
VSSS (1-5)	3.95 (0.51)	3.92 (0.54)
LQL (1-7)	4.55 (0.85)	4.49 (0.90)

At 1-year follow-up, assessments were attempted only for the 166 patients still in contact with the service. Twenty-three patients were too unwell to interview, and 22 had follow-up staff assessment but refused to complete the self-administered instruments, giving a sample of 121 patients (73% of the patients still in contact, comprising 46% of baseline sample) with full baseline and follow-up assessments. Compared with the 121 patients with full data, the 144 patients for whom complete follow-up data were unavailable were more likely to be employed ($t=2.7$, $p<0.01$), have a non-psychotic diagnosis ($t=3.7$, $p<0.01$), and have higher GAF ($t=3.0$, $p<0.01$), lower DAS ($t=2.8$,

p<0.01), lower BPRS (t=2.1, p=0.04), and higher staff CAN-unmet (t=2.4, p=0.02), CAN-met (t=2.0, p=0.05) and patient CAN-unmet (t=10.7, p<0.01) ratings. There was no difference in baseline quality of life.

5.6.1 Hypothesis (i): Cross-sectional association and isolation

At baseline the correlation for all 265 patients of LQL with patient-rated unmet needs was -0.34 (p<0.001). The regression analysis on LQL for all patients is shown in Table 5.4.

Table 5.4: Regression on quality of life for all patients (n=265), using staff and patient assessments of need separately and combined

	Staff CAN only			Patient CAN only			Staff and patient CAN		
	B	**Beta**	**p**	**B**	**Beta**	**p**	**B**	**Beta**	**p**
Age	-0.01	-1.00	0.096	0.00	-0.07	0.200	0.00	-0.09	0.119
Sex	**-0.26**	**-0.15**	**0.005**	-0.15	-0.08	0.100	**-0.20**	**-0.11**	**0.027**
Marital status	0.16	0.09	0.083	0.16	0.09	0.084	0.15	0.09	0.089
Employment status	-0.13	-0.07	0.225	-0.09	-0.05	0.341	-0.14	-0.08	0.161
Psychosis	**0.33**	**0.18**	**0.001**	**0.28**	**0.16**	**0.004**	**0.28**	**0.16**	**0.003**
DAS	**0.23**	**0.25**	**0.004**	**0.21**	**0.23**	**0.005**	**0.27**	**0.29**	**0.001**
BPRS	-0.08	-0.05	0.608	-0.31	-0.17	0.040	-0.16	-0.09	0.311
GAF	0.01	0.13	0.146	0.01	0.13	0.118	0.01	0.11	0.196
VSSS	**0.76**	**0.45**	**<0.001**	**0.73**	**0.43**	**<0.001**	**0.65**	**0.39**	**0.000**
staff CAN-unmet	**-0.16**	**-0.29**	**<0.001**				**-0.13**	**-0.24**	**0.001**
staff CAN-met	-0.04	-0.12	0.068				-0.01	-0.04	0.552
patient CAN-unmet				**-0.09**	**-0.22**	**<0.001**	**-0.08**	**-0.19**	**<0.001**
patient CAN-met				**-0.09**	**-0.22**	**<0.001**	**-0.09**	**-0.21**	**<0.001**

Bold = $p<0.05$

The model comprising solely sociodemographic predictors accounted for 1% of variance in LQL, with diagnosis (psychosis versus non-psychosis) added still accounted for only 1%, and with all variables except CAN added accounted for 30% (not shown). Adding the staff CAN ("Staff CAN only" columns in Table 5.4) gave a model accounting for 34%, adding the patient CAN ("Patient CAN only" columns) gave a model accounting for 37%, and adding both staff and patient CAN ("Staff and patient CAN" columns) gave a model accounting for 40% of the variance in LQL.

For this final model including all variables, higher quality of life was associated with being male, having a diagnosis of psychosis (as opposed to depressive or other neurosis, personality disorder or other or no psychiatric diagnosis), having higher disability, having higher satisfaction with care, having fewer unmet needs (however rated), and fewer patient-rated met needs. For the combined model, the CAN predictors were patient-rated met needs (B=-0.09, 95% CI -0.13 to -0.05), patient-rated unmet needs (B=-0.08, 95% CI -0.12 to -0.04), and staff-rated unmet need (B=-0.13, 95% CI -0.21 to -0.05). These regression coefficients imply that, for example, two people who differed by one patient-rated unmet need would tend to differ by 0.09 LQL units (scale 1 to 7). The same analysis undertaken on the 121 long-term patients (not shown) also found that patient-rated unmet needs were inversely associated with quality of life (B=-0.09, 95% CI -0.15 to -0.03).

Overall, the number of patient-rated unmet needs was cross-sectionally inversely associated with subjective quality of life, whether or not clinical and social variables were controlled. The null hypothesis can be rejected, and hypothesis (i) is confirmed.

5.6.2 Hypothesis (ii): Longitudinal association

The mean staff CAN-unmet rating at 1-year follow-up was 0.99 (a reduction of 0.12), mean staff CAN-met rating was 3.12 (an increase of 0.44), and mean LQL rating was 4.63 (an increase of 0.14). The results of a regression using baseline data on follow-up LQL for the long-term patients are shown in Table 5.5.

Table 5.5: Regression on follow-up quality of life for long-term patients (n=121), using baseline variables with and without baseline quality of life included

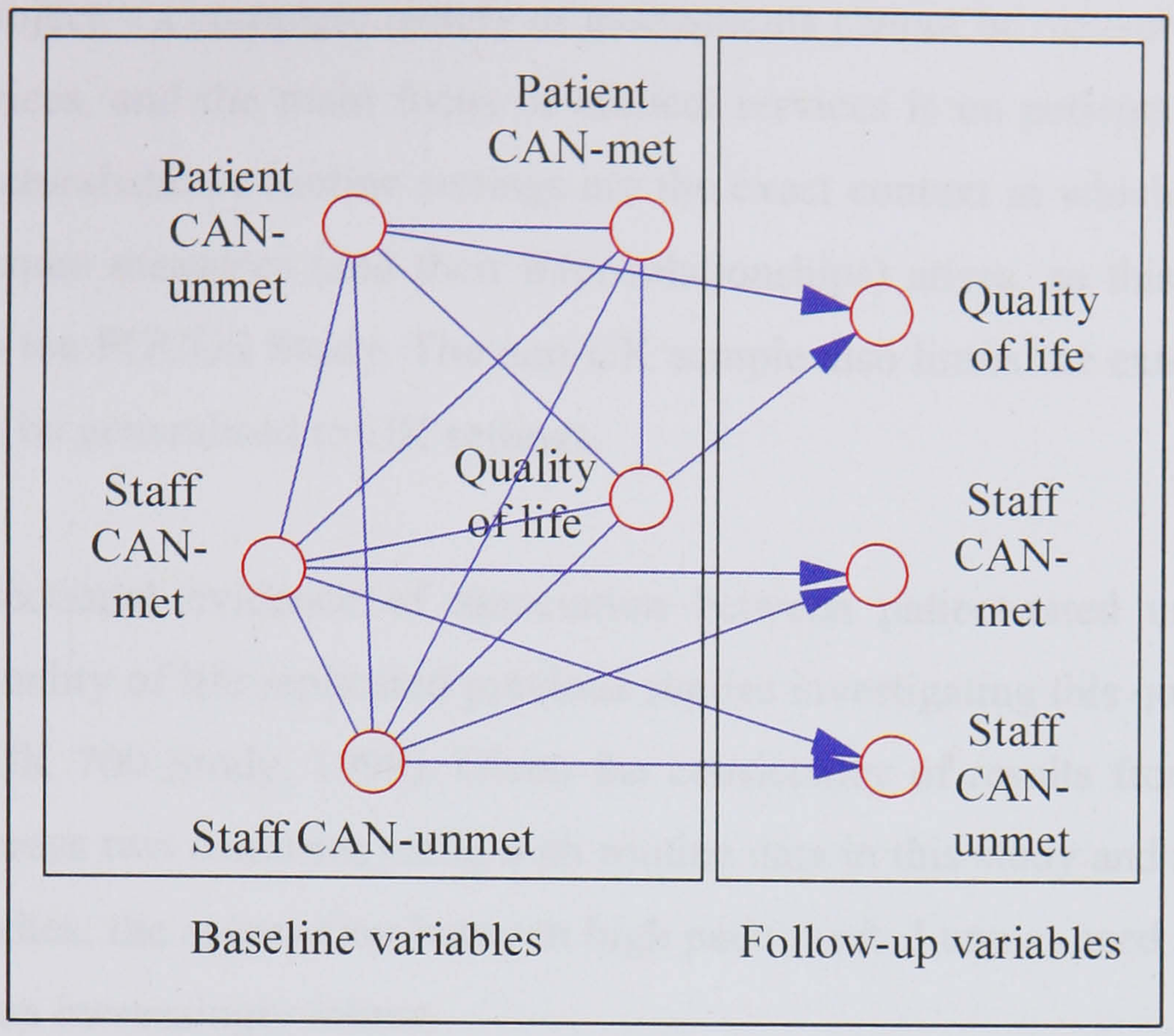
Variable	Excluding baseline LQL			Including baseline LQL		
	B	**Beta**	**p**	**B**	**Beta**	**p**
DAS	0.32	0.34	0.013	0.05	0.05	0.650
BPRS	0.21	0.13	0.327	0.33	0.19	0.062
GAF	**0.02**	**0.33**	**0.012**	0.01	0.16	0.144
VSSS	0.28	0.17	0.053	-0.07	-0.04	0.570
staff CAN-unmet	-0.10	-0.18	0.112	-0.06	-0.11	0.213
staff CAN-met	0.01	0.04	0.669	0.02	0.06	0.449
patient CAN-unmet	**-0.15**	**-0.41**	**<0.001**	**-0.08**	**-0.23**	**0.002**
patient CAN-met	**-0.08**	**-0.19**	**0.033**	-0.02	-0.04	0.585
LQL				**0.62**	**0.63**	**<0.001**

Bold = $p<0.05$

The model accounted for 33% of the variance in quality of life at follow-up with baseline LQL excluded ($r^2 = 0.33$, adjusted $r^2 = 0.28$), and 58% with baseline LQL included ($r^2 = 0.58$, adjusted $r^2 = 0.55$). Apart from baseline LQL, the best predictor (*i.e.* the largest Beta) of follow-up quality of life was baseline patient-rated unmet need (B=-0.08, 95% CI -0.21 to -0.09 excluding baseline LQL, B=-0.08, 95% CI -0.14 to -0.03 including baseline LQL). The null hypothesis can be rejected, and hypothesis (ii) is confirmed.

The relationship between baseline and follow-up values for CAN and LQL was elaborated using graphical chain modelling, shown in Figure 5.1.

Figure 5.1: Graphical chain model of baseline and follow-up needs and quality of life variables (n=121)



When controlling for all possible relationships between variables (a feature of graphical modelling), the follow-up quality of life score was best predicted by baseline patient CAN-unmet and LQL scores. In addition, the follow-up staff CAN assessments were predicted by their baseline levels, and follow-up staff unmet needs were predicted by baseline staff met needs. Follow-up staff met needs were also predicted by baseline staff unmet needs using all data, but this relationship became insignificant when one outlier (identified by analysis of residuals) was omitted.

5.7 Discussion of confirmatory re-analysis

This study investigated whether level of unmet need is temporally associated with level of subjective quality of life. Use of routine outcome data maximises its generalisability to other mental health services. Patient-rated unmet need, and to a lesser extent patient-rated met need, was cross-sectionally associated with subjective quality of life, controlling for other sociodemographic and clinical variables. The level of subjective quality of life at one-year follow-up was also predicted by baseline level of patient-rated unmet need, whether or not baseline subjective quality of life was included (*i.e.* the prediction was not simply due to being a patient measure).

Design limitations include the absence of CAN patient ratings at follow-up, only measuring outcomes at two time points (and so not allowing examination of temporal

precedence in change scores), and not assessing patients out of contact with services at follow-up. These limitations arise from the naturalistic nature of the South Verona Outcome Project – a complete battery of assessments cannot be measured repeatedly in routine services, and the main focus of clinical services is on patients still in contact. However, naturalistic or routine settings are the exact context in which the question or routine outcome measures (and their inter-relationships) arises, so this study has high relevance to the FOCUS Study. The non-UK sample also limits the extent to which the findings can be generalised to UK settings.

The cross-sectional evidence of association between patient-rated unmet need and subjective quality of life replicated previous studies investigating this question (Slade et al, 1999b; UK 700 Study, 1999). Given the consistency of results from three distinct databases across two countries using both routine data in this study and research data in previous studies, the association between high patient-rated unmet need and low quality of life appears increasingly robust.

The graphical chain model shown in Figure 5.1 also demonstrates that there is a partition between staff and patients ratings. This is indicated by the (patient-rated) quality of life being associated only with patient-rated needs and quality of life at baseline, and staff-rated needs at follow-up only being associated with staff-rated needs at baseline. This accords with findings from previous studies that staff and patient ratings of need differ systematically (*e.g.* Slade et al, 1998; Wiersma et al, 1998; Lasalvia et al, 2000; Hansson et al, 2001).

The analysis also provided some evidence that longitudinal association was present – baseline patient-rated unmet need was associated with quality of life one year later, whether or not baseline quality of life was included. This is consistent with the evidence reviewed in Sections 5.2 and 5.3 suggesting that high levels of patient-rated (but not staff-rated) unmet need may, unlike symptomatology, disability or functioning, actually *cause* low levels of subjective quality of life. Of course, it may be that patient-rated unmet need is a mediating or intermediate factor, but not necessarily a causal factor. The stronger test of longitudinal association will be provided when data allows investigation of whether *change* in patient-rated unmet need temporally precedes change in quality of life. This requires measurement at a minimum of three time points, and will be a feature of the FOCUS RCT.

5.8 Recent empirical data

The literature review presented in Sections 5.2 and 5.3 was completed in 2001, and informed the design of the FOCUS RCT. To update the review in the light of more recent publications, a Medline search from January 2001 until October 2003 was undertaken to identify studies with "Camberwell Assessment of Need" or "Needs for Care Assessment Schedule" in their title or abstract. This search identified 47 studies. Reviewing these studies indicated two which were of direct relevance to the relationship between patient-rated unmet need and quality of life.

Foldemo and Bogren (2002) published a 5-year follow-up of patients with schizophrenia who were in-patients for more than 3 months in 1993. 19 patients were interviewed 6 months post-discharge, and 17 were followed up 5 years later. Assessments involved staff and patient ratings of CAN, and patient assessment of quality of life using the Quality of Life Scale (Skantze, Malm, Dencker et al, 1992). Quality of life was higher and patient-rated and staff-rated unmet need were lower at follow-up. Needs and quality of life were not statistically compared.

Hansson and colleagues re-analysed the data from the Nordic multi-centre study of schizophrenia (Hansson et al, 2001) to investigate the cross-sectional relationship between needs and quality of life in 418 people with schizophrenia (Hansson, Sandlund, Bengtsson-Tops et al, 2003). They found that higher patient-rated and staff-rated unmet needs assessed using CAN were associated with lower quality of life ratings on LQL. A forward stepwise regression analysis used age, sex, social network, self-esteem, symptoms, duration of illness and (in the final block) patient-rated met and unmet need. The regression model accounted for 41% of the variance in quality of life, with patient-rated unmet need accounting for 6%. Patient ratings of unmet need in each of the 22 CAN domains were also entered in a new regression model, with five domains contributing to the 31% variance explained by the model: social relationships (10%), accommodation (4%), psychotic symptoms (2%), benefits (1%) and childcare (1%). Only unadjusted r^2 statistics were reported.

Overall, the original review of needs and quality of life, the re-analysis of existing data from South Verona, and the updated review are all consistent with the hypothesis that higher patient-rated unmet need causes a poorer quality of life. This will therefore be

one hypothesis to be tested in the FOCUS RCT, and patient-rated rather than staff-rated unmet need will be the primary clinical outcome.

The FOCUS RCT will involve routine data, and this final design concern is now considered.

5.9 The nature of routine outcome data

Mental health services are intended to be evidence-based, which involves considering efficacy (the ability of the intervention to produce benefit if applied ideally), effectiveness (the benefit that actually occurs when a treatment is used in practice), and efficiency (the resources required to produce a 'unit' of health gain) (Andrews, 1999). The standard view is that the effect size in efficacy studies is diluted when the intervention is applied routinely. For example, Weisz, Weiss and Donenberg (1992) found that the effect of psychotherapy for children declined as the studies increasingly resembled routine practice, with no effectiveness evident for entirely routine settings. Similarly, Shadish and colleagues found that effectiveness evidence was consistent with efficacy evidence in 56 studies of adult clinics, although there was a paucity of outcome data from routine clinical settings (Shadish, Matt, Navarro et al, 1997). This view assumes that the translation of treatments into routine settings merely weakens, rather than changes, the relationship between inputs, processes and (of particular relevance here) outcomes.

Even if the true relationship does not change, the observed relationship may be weaker within routinely-collected data for two statistical reasons: shrinkage and attenuation. These are now considered.

5.9.1 Reduced effect size due to shrinkage

Shrinkage means that the fit of the regression model to the data from one study is likely to reduce when the same variables are measured in another study of the same population and data quality. Possible causes include measurement error and a low participant numbers to predictor variables ratio. This problem can be addressed statistically using the adjusted-r^2 statistic, which adjusts for the degrees of freedom in the model. The formula for this statistic is $1 - ((1 - r^2)\ ((n-1)/(n-k-1)))$, where n is the number of subjects, k is the number of predictor variables and r^2 is the unadjusted estimate of the proportion of variance accounted for by the regression model. The adjusted r^2 statistic

gives a measure of the association between the variables which can be expected to be replicated in other studies where the population and the data quality are similar (Altman, 1991).

For example, in the regression involving all 22 CAN domains reported by Hansson and colleagues (2003) which was described in Section 5.8, the sample size to predictor variable ratio was 418 to 28 (approximately 15 to 1). This indicates shrinkage is likely. The model accounted for 31.3% of the variance in quality of life. Therefore $r^2 = 0.313$, $n = 418$ and $k = 28$. This indicates that the adjusted r^2 statistic is 1-((1-0.313)((418-1)/(418-28-1))) = 0.26. In other words, 26% of the variance is likely to be accounted for in a future replication.

5.9.2 Reduced effect size due to attenuation

The second statistical reason is increased **attenuation** due to measurement error. Routinely collected data will be of a lower quality than researcher-collected data for a number of reasons, including technical proficiency, training and motivation of the clinical rater compared with the research rater, lower return rates, and higher rates of selective attrition. Measurement errors from any source (*e.g.* rater proficiency, reliability of the assessment) are assumed to be random in classical test theory, so this can be expressed statistically as a reduction in the reliability of the measure (irrespective of the reason for error in measurement). To illustrate the effect of attenuation, the percentage of complete agreement between raters for the 22 CAN items ranged from 82% to 100% (Phelan et al, 1995). Assuming from this that the reliability of CANSAS when used in research studies is 0.90, the strength of the association in routine data will be the reliability when used routinely divided by 0.90. If the reliability is reduced to 0.60, then the correction will be (0.6/0.9) = 0.67 – in other words, attenuation will on average reduce the observed level of association by one third.

5.9.3 Different relationships in routine data

However, it may be that different patterns of relationships between outcomes emerge when interventions are applied within routine settings. This has important implications for the way services are planned and organised. Before advocating that more resources be allocated to those with higher levels of patient-rated unmet need, it is important to establish whether the relationship between patient-rated unmet need and quality of life is preserved in practice. For instance, it may be that greater health gain accrues through

allowing a routine service to specialise in working with one diagnostic group, and hence to attain an expertise closer to that found in intervention studies, than to focus efforts on people with higher levels of unmet need from disparate diagnostic groups. This raises the important question of whether relationships between outcome measures are retained when data are collected routinely from busy clinicians and 'typical' patients who are not selected according to stringent inclusion criteria.

The FOCUS RCT will be designed to allow this aspect to be investigated. Routine clinical data will be collected from staff and patients. The staff data will be assessed by the clinician without the involvement of a trained researcher. The patient data will be provided by the patient, again without the involvement of a trained researcher. The patient sample will be as representative as possible of those using adult mental health services, in contrast with efficacy studies which use highly specified inclusion criteria. Therefore the data collected will be relevant to routine clinical settings. In addition to evaluating the effectiveness of the intervention (routine outcome assessment), this will allow investigation of whether the relationship found between patient-rated unmet need and quality of life in research studies is preserved in routinely-collected data.

The relationship between patient-rated unmet need and quality of life has been reviewed, and will inform one of the hypotheses in the FOCUS RCT. The method for this trial is described in Chapter 6.

Chapter 6

Method

6.1 Aims

Chapters 2 and 3 have reviewed theory relevant to routine outcome assessment, leading to the development of the FOCUS Model in Chapter 4. Chapter 5 developed a hypothesis regarding the association between patient-rated unmet need and quality of life. This chapter describes the FOCUS Randomised Controlled Trial (RCT), which had two aims. The first aim was to investigate whether routine collection and feedback of outcomes produced benefits for individual patients, as predicted by the FOCUS Model. The second aim was to investigate whether the relationship found between patient-rated unmet need and quality of life in research studies reviewed in Chapter 5 was replicated in routinely collected data, and to elaborate this relationship through analysis of repeated measure data.

6.2 Hypotheses

The primary hypotheses to be tested were:

1. The routine collection and feedback of outcome information for 7 months will lead to at least 1.0 fewer patient-rated unmet needs, as measured using CANSAS
2. The routine collection and feedback of outcome information for 7 months will lead to an increase of at least 0.25 points in quality of life, as measured using MANSA
3. Baseline level of patient-rated unmet need will predict follow-up level of quality of life.

The UK700 Study found that complex social programme interventions have differential effects for patients on the basis of their premorbid IQ (Hassiotis, Ukoumunne, Byford et al, 2001). This raises the question of whether premorbid IQ will influence the effectiveness of routine outcome assessment. In addition, an important question for clinical practice is whether there is any evidence of differential effectiveness of the intervention by staff profession. Therefore two sub-group analyses were planned: premorbid IQ and professional group.

6.3 Design

This exploratory randomised controlled trial investigated the impact of routine outcome assessment. Clinical trials methodology was used to evaluate the effectiveness of routine collection and feedback of carefully chosen process and outcome measures.

The RCT involved two groups: the control group who received treatment as usual, and the intervention group who received treatment as usual plus routine outcome assessment. The control and intervention groups are described in detail in Section 6.14. An RCT design was chosen to maximise the scientific value of the trial. To enhance the generalisability of any findings, the intervention was tested in an epidemiologically representative sample of adult mental health patients. The unit of randomisation and the level of analysis were individual patients.

6.4 Management, monitoring and registration

Data were collected by MS, two full-time researchers, one temporary researcher (for one month), and two visiting psychiatrists from Verona, Italy. One of the full-time researchers was employed for 24 months from February 2001 (although the post-holder resigned in December 2001, and was replaced in January 2002), and the other for 26 months from August 2001. Both psychiatrists visited for six months, one in 2002 and one in 2003. All data collection was carried out under the supervision of MS, who was responsible for all aspects of the running of the trial. The RCT was conducted in accordance with the MRC Guidelines for Good Clinical Practice in Clinical Trials (1998), and relevant legislation including the Data Protection Act (1998) and the Research Governance Framework for Health and Social Care (2001).

The RCT was granted ethical approval by the Institute of Psychiatry Local Research Ethics Committee (Ref 012/00), and was overseen by a Trial Steering Committee (TSC) chaired by Prof Tom Burns (Department of Psychiatry, Oxford University). The TSC included the functions of a Data Monitoring and Ethics Committee, and the independent trial statistician was Ian White. Although no formal stopping rules were developed, interim analysis of adverse events was undertaken at each TSC meeting (every 6 to 9 months). The RCT was assigned the International Standard Randomised Controlled Trial Number (ISRCTN) of 16971059. The trial was registered on the Current Controlled Trials *meta*Register (reference 16971059), and on the National Research Register (Publication ID N0042063032).

6.5 Sample size

The power analysis for the exploratory trial involved an estimation of effect size. One product of the exploratory RCT was intended to be the identification of criteria for a more refined power analysis to inform a future definitive RCT.

Hypothesis 3 (unmet need causes quality of life) will be tested by pooled analysis of baseline and follow-up data from both groups, and by analysing the repeated measures data from the intervention group. For the latter analysis, with 85 patients a correlation of 0.3 (*i.e.* a cross-sectional association where one variable accounts for 10% of the variance in another) with $p<0.05$ will be detected with a power of 0.80.

The control group will be compared with the intervention group to test hypothesis 1 (routine outcome assessment reduces unmet need) and hypothesis 2 (routine outcome assessment improves quality of life). The CANSAS-P unmet needs measure is assumed (on the basis of unpublished data from the PRiSM Psychosis Study) to have a standard deviation of 1.7 and a correlation from t1 to t2 of 0.32. Assuming $p<0.05$ and analysis of covariance is used to compare t2 values whilst adjusting for t1 levels, an intervention group of 85 will require a control group of 50 to detect a change of 1.0 patient-rated unmet needs with a power of 0.94. The MANSA was assumed to have a standard deviation of 0.5 (Priebe et al, 1999) and a correlation from t1 to t2 of 0.5 (unpublished PRiSM Psychosis Study data), so with the same assumptions will detect a change of 0.25 in quality of life rating with a power of 0.9.

The intervention group will therefore need 85 patients, and the control group 50 patients. More patients are needed in the intervention group than the control group to ensure the sample size for testing hypothesis 3 (unmet need causes quality of life) is adequate. Dropouts in this study will comprise patients who die, move, or are discharged from mental health services. Initially it was assumed that the drop-out rate would be 20% with data unobtainable for a further 5%. The initial target was therefore that the intervention group would comprise 113 patients, and the control group would comprise 67 patients, with a total sample of 180 patients. Due to higher-than-anticipated retention rates, this target sample size was subsequently reduced to 160 with no loss of power – as described in Section 6.16.

6.6 Sample setting and location

The sample was taken from the caseloads of all eight community mental health teams (CMHTs) in the London Borough of Croydon, South London. Croydon has a population of 319,000, and an average Mental Illness Needs Index (Glover, Robin, Emami et al, 1998) score of 100.1, with a range from 81.7 (most affluent electoral ward) to 111.1 (most deprived electoral ward). This indicates that the level of social deprivation varies widely within what is overall an area with average levels of deprivation, making it a highly nationally representative location. There are approximately 3,500 patients under the care of mental health services in Croydon.

Croydon adult mental health services are organised into three localities (North, Central and South), containing eight CMHTs based in four resource centres. Three CMHTs are in one resource centre in North locality, three CMHTs are in one resource centre in Central locality, and two CMHTs are in separate resource centres in the less densely populated South locality. An information system – the Patient Administration System (PAS) – is used throughout mental health services in Croydon, which records basic sociodemographic measures, clinical diagnosis, and service contact information.

At the time of the study there was no routine collection, recording or use of standardised assessment or outcome measures in any of the eight Croydon teams. During the study the routine use of HoNOS and a standardised risk assessment was introduced in other parts of the mental health trust, but not in Croydon.

6.7 Eligibility criteria

Patients were included who met all three of the following criteria:

1. The patient was on the caseload of an adult mental team in Croydon on 1 May 2001
2. The patient has been on the caseload for at least 3 months
3. The patient was aged between 18 and 64 inclusive.

The goal of inclusion criterion 2 was to exclude people who were referred for one-off assessments, such as for management advice (*e.g.* medication levels), benefits eligibility, or for a letter to the Housing Department.

6.8 Sample selection

To ensure that the sample selected was representative of the overall group of patients, stratified and pure random sampling were considered. Stratified random sampling identifies all patients fitting each cell ('stratum') defined by identified prognostic factors, and then randomly selects within each sub-group of identified patients. A **prognostic factor** is a characteristic thought likely to correlate with the patient's subsequent response to treatment (Everitt and Wessely, 2004). This ensures that the selected sample has approximately the same proportions of each prognostic factor in each stratum as in the whole group. Pure random sampling has the benefit over stratified random sampling of being more transparently random in its sample selection, but where there is a relatively small sample size it does not guarantee that the final sample is representative in terms of the specific prognostic factors chosen to define the strata. Given the relatively small sample size, stratified sampling was used. The sample selection was performed independently by Dr Morven Leese (Statistician).

Prognostic factors in this context are those for which there is evidence that they impact directly on content of care or indirectly on process of care. Age impacts on amount of mental health care, with older people receiving less care (Horwitz & Uttaro, 1998). Gender impacts on patterns of service use (Prior, 1999), and ethnicity impacts on type of treatment offered (Nazroo, 1999). Clinical diagnosis, in particular whether a functional psychosis or not, impacts on both content and process of care. CMHT impacts on both process and content of care, since there are different services (*e.g.* availability of psychological and social interventions) and cultures within the different CMHTs. Therefore prognostic factors identified were age, gender, ethnicity, clinical diagnosis and CMHT.

Which of these prognostic factors were available in the PAS? Date of birth and gender were recorded. Ethnicity was only available for about 20% of patients. Clinical diagnosis was recorded but anecdotally known to be unreliable, due to the lack of standardised training and being based on initial presentation rather than updated over time. The CMHT was not recorded but Consultant (Psychiatrist) was a good proxy measure, since all patients had their care listed under individual Consultants, and each Consultant had worked or was working with one CMHT. Therefore the prognostic factors available from the PAS for sample selection were age, gender and CMHT.

Information was obtained from the PAS about all patients in contact with adult mental health services in Croydon on 1 May 2001. There were 1182 patients listed under adult mental health service consultants, and who were aged between 18 and 64. This was termed the **base population**.

The sample was then selected to ensure that age, gender and CMHT were in proportion to their representation in the overall sample. This approach of equal sample fractions allows straightforward analysis without weighting, since the sample is fully representative of the base population. Men and women were equally represented in the base population, but had different age distributions, with more women in the highest tertile. The sample was therefore stratified by tertiles within gender, to give 6 age-gender combinations having approximately equal proportions in the base population, as shown in Table 6.1.

Table 6.1: Sample selection by age and gender

		Base population		Sample	
Stratum	**Age range**	**n**	**%**	**n**	**%**
Men 1	18.41-31.70	199	16.8	30	16.6
Men 2	31.73-41.96	200	16.9	30	16.6
Men 3	42.97-64.88	200	16.9	30	16.6
Women 1	19.11-35.09	194	16.4	30	16.6
Women 2	35.16-47.05	194	16.4	30	16.6
Women 3	47.16-64.92	195	16.5	30	16.6
Total		**1182**	**100**	**180**	**100**

The distribution of patients across the 8 CMHTs was not equal, so a proportionately higher recruitment target was set for CMHTs dealing with more patients than average, as shown in Table 6.2:

Table 6.2: Sample selection by Community Mental Health Team

CMHT	Base population		Sample		
	n	%	n	%	Number from each stratum
1	154	13	24	13	4
2	192	16	30	17	5
3	153	13	24	13	4
4	170	14	24	13	4
5	173	14	24	13	4
6	122	10	18	10	3
7	143	12	18	10	3
8	75	6	18	10	3
Total	**1182**	**100**	**180**	**100**	

Each patient was allocated a random number generated by SPSS Version 10. The complete list of 1182 patients was then sorted by CMHT, stratum, and random number. The first set of patients from each stratum in each CMHT formed the initial sample.

After 1 May 2001 the casenotes of patients were reviewed. Since time in contact with services was not reliably accessible from the PAS, eligibility criteria of (i) being in contact, (ii) having been in contact for the previous 3 months and (iii) being aged between 18 and 64 inclusive (all with reference to 1 May 2001) were re-evaluated. Where the patient was found not to be eligible for inclusion (*e.g.* because they had not been in contact with the service for 3 months on 1 May 2001), they were substituted with the next person in the appropriate CMHT / stratum combination.

In practice, the PAS data proved unreliable, with many patients not on the list whom staff identified as on their caseload, and many listed patients unknown or discharged. Whereas the PAS identified 1,182 patients, service managers believed that approximately 3,500 patients were attending CMHTs. Therefore a protocol deviation in the sample selection method was agreed with the Trial Steering Committee in November 2001 (at which point 48 patients had been recruited). The new method was to identify with each member of staff how many patients they were prepared to have in the study. The sample for that member of staff was then selected from the patients who

were identified for that CMHT using the stratified random sampling procedure – they were given the intended sample for their CMHT and asked to identify any patients on their caseload. When this was insufficient, the remaining list of the PAS patients (who were identified from the PAS but not randomly selected) was reviewed, to identify any on their caseload. When both those sources were exhausted, the CMHT member was asked to identify a further 10 patients from their caseload, and patients were chosen from the list by the researcher using a random number table approach (form shown in Appendix 1).

6.9 Allocation

Information was collected during the baseline assessment which allowed a characterisation of the sample in terms of the five prognostic factors:

- Age (tertiles, as described above)
- Gender (male or female)
- Ethnicity (white or non-white)
- Clinical diagnosis (functional psychosis or other)
- CMHT (1 to 8).

Following baseline assessment, patients were randomly allocated into the intervention group (n=113) or the control group (n=67). Random allocation was used to ensure the allocation was not influenced by the baseline assessment or clinician preference, and that any unmeasured prognostic factors would be balanced.

In the random allocation it was important to ensure balance on known prognostic factors, which can be done in one of three ways: simple random allocation, stratification and minimisation. The first option was not appropriate due to the relatively small sample size, which did not guarantee balance between the two groups. Stratification would have required 3 x 2 x 2 x 2 x 8 = 192 strata, which would not be feasible with 180 patients. Therefore the restricted randomisation approach of minimisation was used to ensure balance between the intervention and control groups in the five prognostic factors.

Minimisation, or adaptive randomisation procedure, is a randomisation approach in which the chance of allocation to a particular treatment is adjusted to account for any

existing imbalances in the baseline characteristics of the sample (Everitt & Wessely, 2004). The aim is to minimise imbalances in the distribution of known prognostic factors. Minimisation has a known problem of 'yo-yoing' – when several consecutive participants who are indistinguishable in terms of prognostic factors are allocated, and the already-allocated people in the two arms are already balanced on the prognostic factors, then the resulting allocation sequence will be partly predictable. The potential for this form of allocation bias was reduced by the use of five balancing factors with 192 possible permutations.

An inadequate allocation method has been found to be associated with an exaggeration of odds ratios by 41%, and an inadequately specified allocation method with an exaggeration by 30% (Schulz, Chalmers, Hayes et al, 1995). Therefore the random allocation was performed by Ian White (MRC Biostatistics Unit, Cambridge University), who was blind to the results of the baseline assessment other than the prognostic factors needed for the random allocation.

The characteristics of each batch of patients whose baseline assessment had been completed in the previous week were sent by email as an Excel attachment, with patients listed only by an identification number. The minimisation was undertaken using a purpose-written Stata program. The team variable was given twice the weight of the other factors, since balance within teams was important to ensure the spread of the study across Croydon. Randomisation was in the ratio 113 : 67 in favour of the intervention group. The reason the allocation was not in the ratio 1 : 1 was explained in Section 6.5. The allocations for each identification were returned by email. To summarise, participants were enrolled by the researcher, following which participants were randomly allocated by the external trial statistician using minimisation.

6.10 Measures

Two sets of outcome measure were used: one set as part of the intervention, and another set to evaluate the impact of the intervention. The first set will be called the **postal questionnaire**, and the second set the **evaluation measures**. The data from the postal questionnaires were also used to evaluate the intervention, and indeed contained the two primary outcome measures (MANSA and CANSAS). The postal questionnaire and the evaluation measures are shown in Appendix 1.

6.10.1 Postal questionnaire

As discussed in Chapter 4, TAG, MANSA, HAS, and CANSAS were chosen as the outcome measures for routine completion in the intervention group.

The TAG is a 7-item assessment of the severity of a person's mental health problems (Slade et al, 2000). It was developed through the use of innovative consensus workshops and a Delphi Consultation to obtain consensus about identifying the 'severely mentally ill' – the priority group for mental health services. It is completed by making one tick to indicate level of severity in each of 7 domains: (i) intentional self-harm; (ii) unintentional self-harm; (iii) risk from others; (iv) risk to others; (v) survival needs / disabilities; (vi) psychological needs / disabilities; and (vii) social needs / disabilities. The scale is "None", "Mild", "Moderate" and "Severe" (4-point scale) for domains (ii), (iii), (vi) and (vii), with an extra "Very Severe" domain possible for the remaining 3 domains (which may require immediate action). The time frame for the TAG is the previous month. The TAG is intended for use without training by any health professional. The summary TAG score (range 0 to 24, low score better) is calculated by summing the columns scores, with 0 for None up to 4 for Very Severe, and gives a global measure of severity.

The MANSA is derived from the Lancashire Quality of Life Profile (LQL) (Oliver, 1991), which has been widely used in Europe as a quality of life measure (Oliver et al, 1997). The LQL has a number of shortcomings for routine use, including its administration time (30 minutes), the inclusion of non-quality of life scales (such as the psychopathology measured by the affect balance scale), non-discriminatory and hence redundant items, inconsistent language, and no question on sexual life. To remedy these shortcomings, the MANSA was developed, which has a high correlation (item range 0.83-0.99) with LQL scores (Priebe et al, 1999). MANSA comprises 16 items – four "objective" yes/no questions (having a friend, seeing a friend, being accused of a crime, and being a victim of physical violence) and twelve subjective questions measuring satisfaction with life as a whole, job (or sheltered employment, or training/education, or unemployment/retirement), financial situation, number and quality of friendships, leisure activities, accommodation, personal safety, people that the patient lives with (or living alone), sex life, relationship with family, physical health, and mental health. The time frame for the subjective questions is not specified, so was taken as cross-sectional. Only the subjective items were assessed in the postal questionnaire, since the goal was

to identify subjective well-being. The objective questions were only asked at baseline and follow-up, using the Supplementary-P form. Each item is rated on a 7-point satisfaction scale, from 1 = "Couldn't be worse" to 7 = "Couldn't be better". Although tested with an interviewer, the scale was amended for self-administration in this study – the implications of this change are explored in Section 8.5.2. The summary score is the mean of the 12 subjective items (range 1 to 7, high score better).

The HAS is a modified version of the Helping Alliance measure described by Priebe and Gruyters (1993). The original questionnaire involved an interviewer asking the patient five questions: a) Do you feel understood by your case manager? b) Do you feel criticized by your case manager? c) How much is your case manager committed to and actively involved in your treatment? d) Is the treatment you are currently receiving right for you? e) How do you feel immediately after a session with your case manager? The first four questions were rated on a 100-mm long visual analogue scale, with each 10-mm interval marked, from 0 = "Entirely" to 10 = "Not at all". The fifth question allowed responses of either "Better" or "Unchanged / Worse". In the Priebe and Gruyters study, all of questions d) and e) and the sum score over all 5 items predicted hospitalisation rates over 20 months.

The amended patient version (**HAS-P**) separates "Unchanged" and "Worse" for the final question, with "Worse" coded as 0, "Unchanged" as 5 and "Better" as 10 (McCabe, Roeder-Wanner, Hoffmann et al, 1999). It also adds a question regarding trust in and competence of the therapist, and is self-rated by the patient. A staff version (**HAS-S**) has also been developed with equivalent questions a) and c), and three new questions: i) Do you get along with the service user? ii) Do you look forward to meeting the service user? iii) Do you feel you can help the service user and treat him/her effectively? Two further qualitative questions ask for positive and negative aspects of the relationship with the patient. Hence both HAS-P and HAS-S have 5 questions rated on an 11-point scale, and in addition HAS-P has one categorical question and HAS-S has 2 qualitative questions. The time frame for HAS was not specified, so was taken as cross-sectional. The complete HAS-S and HAS-P were administered at baseline and follow-up, but the HAS-S questions on qualitative aspects were only assessed at baseline and follow-up (using the Supplementary-S form). The psychometric properties of the HAS-S and HAS-P are currently being investigated, and although not ideal, the HAS was chosen because of its feasibility and closeness to a measure with

demonstrated psychometric properties. Relevant scores are reversed so that a high score is good. The HAS-P summary score is the mean of the 6 HAS-P items, and the HAS-S summary score is the mean of the 5 HAS-S items (range for both 0 to 10, high score better).

The Camberwell Assessment of Need Short Appraisal Schedule (CANSAS) (Slade et al, 1999c) is a modified version of the Camberwell Assessment of Need (CAN) (Phelan et al, 1995), which has been widely used across Europe as a measure of need (McCrone, Leese, Thornicroft et al, 2000). The CANSAS is explicitly intended for routine clinical use, and assesses level of need in 22 domains of health and social need. The domains assessed are accommodation, food, looking after the home, self care, daytime activities, physical health, psychotic symptoms, information (about condition and treatment), psychological distress, safety to self, safety to others, alcohol, drugs, company, intimate relationships, sexual expression, childcare, basic education, telephone, transport, money and benefits. For each domain, the possible ratings are 2 = unmet need (current serious problem, regardless of any help received), 1 = met need (no / moderate problem due to help given), 0 = no need or 9 = not known. The time frame for the CANSAS is the previous month. The CANSAS is completed separately by patients (**CANSAS-P**) and staff (**CANSAS-S**). Two summary measures are possible, each with a maximum score of 22 and separately scored from CANSAS-S and CANSAS-P: the summary **met need** is the number of domains with a met need, and the summary **unmet need** is the number of domains with an unmet need (range for both 0-22, low score better).

The staff-completed postal questionnaire measures were TAG, CANSAS-S and HAS-S. The patient-completed postal questionnaire measures were CANSAS-P, MANSA and HAS-P. The primary end-points for the trial are patient-rated unmet need (unmet need from CANSAS-P) and quality of life (from MANSA).

6.10.2 Evaluation measures

All measures completed as part of the intervention were assessed at baseline and follow-up, as well as the objective questions from MANSA (**Supplementary-P** form) and the qualitative questions from HAS (**Supplementary-S** form).

Since the baseline assessment also involved explaining the study, obtaining informed consent, and going over the assessments to be completed, a minimal set of

supplementary evaluation measures were collected (so as to reduce burden and maximise subsequent involvement). A conceptual framework that was used to inform the decision about supplemental measures was the World Health Organization categorisation of impairment, disability and handicap (1980). The CANSAS and MANSA measure the actual problems and resulting subjective well-being the patient experiences in their everyday life, and are relatively comprehensive measures of handicap. The TAG is a less direct and comprehensive measure of impairment (which in a mental health context refers to symptomatology) and social disability, so supplemental measures for these two aspects were chosen. The interviewer-rated Brief Psychiatric Rating Scale (**BPRS**) (Overall & Gorham, 1988) was used to assess symptomatology over the previous two weeks – in the FOCUS RCT the 18-item version was used, unlike the 24-item version reported in Chapter 5. The staff-rated Health of the Nation Outcome Scale (**HoNOS**) (Wing et al, 1998) was used to assess social disability over the previous two weeks.

Intellectual functioning was measured at baseline only using the second edition of the interviewer-rated National Adult Reading Test (**NART**) (Nelson, 1982). The NART error score can range from 0 to 50, and can be converted into an estimate of premorbid Full-scale IQ. A NART error score of 25 indicates a premorbid full-scale IQ of 100.

In addition to the standardised assessments used, three further forms were developed. The **Sociodemographic form** was rated at baseline by the researcher from casenotes, and recorded date of birth, gender, ethnicity, education and clinical diagnosis. Categorisations of ethnicity and educational level were taken from the 1991 census. The **Casenote form** was rated at baseline and follow-up by the researcher from casenotes, and recorded the number and range of care plan entries, and current living arrangements. The form was based on a locally-developed but unstandardised taxonomy of care plan categories. The information on both the sociodemographic and the casenote form was verified with staff, and where necessary supplemental information from staff or patient was incorporated.

If no difference in outcome was found between the control and the intervention group, then it would be important to try to identify the point(s) where the FOCUS Model (described in Figures 4.1 and 4.2) broke down. For example, did the staff and patient read the feedback? Did staff plan to offer new care but not actually behave any

differently? Was new care offered, but this made no difference to outcome? Intermediate 'markers' of the proposed links in the FOCUS Model were needed. The goal was to investigate the mechanisms of action, by generating information to inform a future definitive RCT (in the event of a positive finding) or identifying aspects of the FOCUS Model requiring re-consideration or specific testing (in the event of a negative finding). To meet this need, the **Impact of Involvement form** was completed at follow-up only by staff and patient. It assessed whether the feedback was received and read, whether either the process or content of care was reflected on after completing the postal questionnaire or after the feedback, and whether feedback led to discussion of the content or process of care. The three non-standardised forms – Sociodemographic, Casenotes and Impact of Involvement – are shown in Appendix 1.

The points of administration and sources of data are shown in Table 6.3.

Table 6.3: Point of administration and source of data for all assessments

	Baseline			Intervention		Follow-up		
	Casenotes	**Staff**	**Patient**	**Staff**	**Patient**	**Casenotes**	**Staff**	**Patient**
Postal questionnaire								
HAS-S		x		x			x	
HAS-P			x		x			x
CANSAS-S		x		x			x	
CANSAS-P			x		x			x
MANSA			x		x			x
TAG		x		x			x	
Evaluation measures								
Sociodemographic form	x							
Casenote form	x					x		
Supplementary-S		x					x	
Supplementary-P			x					x
NART			x					
BPRS			x					x
HoNOS		x					x	
Impact of Involvement form							x	x

6.11 Methods to enhance data quality

All researchers were trained in standardised assessments through (i) role play with MS; (ii) rating of vignettes; and (iii) observation of assessments. Once training was completed, assessment quality was monitored by double rating 12 patient assessments (11 baseline and 1 follow-up), showing acceptable concordance – 8 (2.8%) of 286 CAN ratings differed, mean difference of 0.14 (scale 0 to 7) in 216 BPRS ratings, and a mean difference of 1.7 (scale range 0 to 50) in 12 NART error score ratings mainly due to one outlier. Further training was provided wherever differences occurred, and the rating by MS was used in the analysis. Once training was completed, the researchers independently arranged and undertook interviews.

6.12 Pilot study

To test the assessments and develop the procedure, a pilot phase involved using the assessments with a convenience sample. The postal questionnaires were completed by two patients, the full patient baseline assessment (postal questionnaire plus evaluation measures) with three patients, and the full staff baseline assessment (postal questionnaire plus evaluation measures) with two staff. After each practice, a pilot questionnaire was administered, comprising 10 questions:

1. What in the questionnaire have you found easy?
2. What in the questionnaire has been difficult?
3. Is there anything in the questionnaire you would change?
4. How did you find the instructions?
5. Were there any questions you wouldn't want to answer - why?
6. Did you understand every bit of the questionnaire? If not, what would make it easier to understand?
7. How would you feel if you were asked to complete this questionnaire every month?
8. What kind of feedback from the questionnaire would you like?
9. How would you feel about your key worker completing a similar questionnaire about your working relationship, needs and mental health?
10. How would you feel about your key worker and yourself receiving the same feedback about your answers on this questionnaire and their answers on a similar questionnaire?

Options for presenting the feedback were also discussed. Patients who participated received a £5 postal order. The assessments were also discussed with a small convenience sample of experienced clinical researchers and patient representatives.

The results of the pilot phase indicated a number of changes:

- The order of presentation of measures in the postal questionnaire was changed so as to finish with HAS, since patients reported it was easier to start by thinking about themselves, and then move to thinking about their relationship with staff.
- Instructions were added to each page of the postal questionnaire, identifying what the page was assessing.
- The option of not completing specific questions was made more explicit.
- The CANSAS was found to be difficult to complete with numerical ratings for categorical variables (*e.g.* 2 for an unmet need). Therefore a new self-rated version was developed, which differed from the original CANSAS in being completed with ticks rather than numbers (with a column for each need rating), and using different colours for each column. The implications of this amendment are discussed in Section 8.5.2.
- For the feedback, there was a staff preference for text-based presentation, and a patient preference for graphic feedback. One option considered was to give both staff and patient a set of options from which to select, including text versus graphical feedback, summary score versus item-specific score feedback, and (for graphical feedback) bar versus line charts. However, this option was discounted because the feedback was intended to be identical to staff and patients, and because this level of individualisation would make it difficult to automate the feedback generation process. Therefore a compromise was implemented, with summary scores over time represented graphically, and item-specific feedback on the most recently completed postal questionnaire presented in text-based form. A demonstration feedback form is shown in Appendix 1.

6.13 Procedure

The full involvement of CMHTs was facilitated in a number of ways, following the Green and Eriksen model of practice change (1988). They propose a three-stage model of predisposing, enabling and reinforcing clinical behaviour change. **Predisposing** involves alerting clinicians to a problem and possible solutions. **Enabling** involves

identifying and removing barriers to change. **Reinforcing** involves maintaining changes through rewards, once they are in place.

6.13.1 Predisposing to change

Several predisposing activities were undertaken:

- A Local Implementation Group (LIG) was formed, comprising local psychiatrists, clinical psychologists, psychiatric nurses, care managers and occupational therapists, along with two mental health service user representatives, the clinical director for Croydon Adult Mental Health Services, the Croydon Health Authority mental health lead, and the Borough Trust Manager. The LIG met throughout the study, to advise on local issues and to facilitate local ownership.
- Posters publicising the study were distributed though local service user networks (with the help of Croydon Mind and the Croydon Advocacy Service), and displayed at each resource centre in staff areas.
- A FOCUS lunch was held to launch the study in each of the 4 resource centres. These lunches were attended by 21 (resource centre for CMHTs 1 to 3), 25 (CMHTs 4 to 6), 20 (CMHT 7) and 12 (CMHT 8) staff. At the lunches the intention that the study be an aid to reflective practice was emphasised. The vision for the study was also presented, comprising 3 elements:
  1. FOCUS finishes on time and to specification
  2. Croydon staff enjoy taking part in the study
  3. The FOCUS team feel valued in their jobs.

  It was intended that the explicit inclusion of Croydon staff within this vision would counter-balance previous experiences of research, which staff reported as being non-inclusive and "hit-and-run", with little feedback of results.
- The FOCUS RCT was presented at the 2001 Croydon Research Forum.
- MS had worked clinically in Croydon, and this link with many participating staff was emphasised
- All consultant psychiatrists were individually written to about the study.

6.13.2 Enabling and reinforcing change

The initial enabling of participation was straightforward – there were no staff refusals to participate when initially approached for baseline assessments. To maintain the practice change, the research team maintained a high visibility within participating CMHTs –

this led to the perception that the study was partly 'owned' by Croydon services. Low-value gifts were distributed during the study, including the supply of biscuits for each CMHT when recruiting. At Christmas individually signed Christmas cards were sent to all staff participants, and a £10 postal order gift was offered to reception and administrative staff. The study was publicised in health and social services newsletters in 2001 and 2002. During the study two further rounds of FOCUS lunches were held. At the end of the study a final FOCUS lunch was held at each of the four resource centres, in which data on the baseline characteristics of study participants from each resource centre were presented, and staff who had four or more patients in the study were named and rewarded with postal orders ranging in value from £5 to £20.

Patient participation was also maximised in several ways. Care was taken at the start of the study to discuss with the patient what to do if they did not respond, and these suggestions (*e.g.* contacting an informal carer) were acted on where appropriate. The research team emphasised their availability to discuss the study with patients when wanted. All participating patients were individually invited to the FOCUS lunches. Finally, a payment of £5 was made for each round of postal questionnaires, including the baseline and follow-up assessments. An incentive payment was used because a majority (13 out of 19) of studies identified in a systematic review showed a positive effect of incentives on response rate, and it was included with the postal questionnaire (rather than as payment on completion) because enclosed incentives were found to be more powerful than promised incentives in enhancing response rates (McColl, Jacoby, Thomas et al, 2001).

6.13.3 Recruitment

The recruitment process for each patient on the PAS sample selection list started with casenote review. Where possible, all relevant casenotes were accessed, including separate notes held by day care services, psychology and care management. The purpose was to check eligibility criteria, to complete the casenote form, and to identify the relevant staff member. The staff informant was identified by surveying recorded contact in the previous 3 months (from when the casenotes were reviewed, not from 1 May 2001), identifying the most frequently-contacted member of staff, and then verifying that there were no recorded plans to discharge the patient in the next month from their caseload (*e.g.* due to the end of a course of therapy). Where discharge plans were recorded, the next most frequent member of staff was identified. This decision was

made in liaison with the mental health service staff, whose advice was also considered regarding the most appropriate staff informant, especially where the casenotes were unclear. The identified person was deemed to be the **staff** informant for the patient.

Once identified, the member of staff was contacted to arrange the baseline staff meeting, and to ascertain whether the patient was well enough to be interviewed. If they were well enough then the patient was contacted by the researcher, and a meeting was arranged as close in time to the staff meeting as possible.

At the patient meeting, verbal and written information about the study was given, which emphasised that all information on the postal questionnaire would be fed back to staff. If written consent to participate was given, the patient was then trained in completing the postal questionnaire, using backward chaining. The first few questions of the CANSAS were completed by the researcher asking questions, giving a verbal description of which rating is appropriate, giving a visual prompt by pointing at the correct box to tick, and then ticking the box. These prompts were then withdrawn in order, so that initially the patient was given the pen but the questions and verbal and visual prompts were given so the patient only had to tick the box, then the visual prompt was withdrawn, then the question was asked but no verbal prompting was given, and finally no questions were asked. A similar approach was used for MANSA and HAS-P where necessary. Administration time for the postal questionnaire was recorded. The evaluation measures were then administered by the researcher, and a demonstration feedback form was provided and explained. Administrative details were confirmed, such as where to send the postal questionnaires and resulting feedback, and who the patient could contact if assistance in completing them was necessary. Where patients were too unwell to complete assessments, only the staff assessment was initially made, and the patient's clinical status was re-evaluated prior to the next round of assessments.

The staff baseline meeting involved an explanation of the study, including emphasising that all staff-rated measures would be fed back to the patient. If written consent to participate was given, the staff postal questionnaire and evaluation measures were then administered. Administration time for the postal questionnaire was recorded.

Once the staff and patient baseline assessment and the Casenote form were completed, the patient was randomly allocated to the intervention or the control group. The

allocation was made within five working days. A letter was sent to the patient and staff giving the result of the random allocation. The control group received treatment as usual, and the intervention group received treatment as usual plus the intervention. Both are now described.

6.14 Description of intervention and control

The **control** intervention was treatment-as-usual. This involved mental health care from the multidisciplinary community mental health team, comprising (when fully staffed) one consultant psychiatrist, one to two junior psychiatrists, one clinical / counselling psychologist, three to four community psychiatric nurses, and sessional input from care managers and occupational therapists as needed.

The **intervention** group received the same care as the control group, and in addition patients and staff (i) were asked to complete a postal questionnaire on a monthly basis; and (ii) were both provided with identical feedback by post at 3-monthly intervals. The postal questionnaire element of the baseline data was used for the round 1 information. One month after allocation, the first postal questionnaire was sent (round 2). Postal questionnaires were then sent monthly. Feedback was sent two weeks after the round 3 and the round 6 questionnaires were sent. The intervention lasted for 6 months, comprising two cycles of postal questionnaire, postal questionnaire, postal questionnaire and feedback. This is described diagrammatically in Figure 6.1 in Section 6.15.

The five-page postal questionnaire sent monthly to patients comprised a cover sheet, CANSAS-P (one page), MANSA (two pages) and HAS-P (one page), along with a prepaid pre-addressed envelope and a £5 postal order. The four-page postal questionnaire for staff comprised a cover sheet, TAG (one page), CANSAS-S (one page) and HAS-S (one page), along with a pre-addressed internal envelope. Staff were asked to return the questionnaire blank if they had not seen the patient in the previous month and had received no new information about them.

Feedback presented summarised information from the staff and patient assessments. Identical feedback information was sent to both staff and patient. The feedback used blue to indicate staff responses and red to indicate patient responses, and comprised graphical and text-based elements. The graphical charts compared summary ratings over

the preceding three months, and the text described item-level responses for the most recent rating.

6.15 Follow-up

Follow-up assessments were made 6 weeks after the second feedback was sent, to allow time for the feedback to impact on care. Follow-up assessment was therefore 7 months after baseline assessment. Casenotes, patients and staff were re-assessed at follow-up by the researcher.

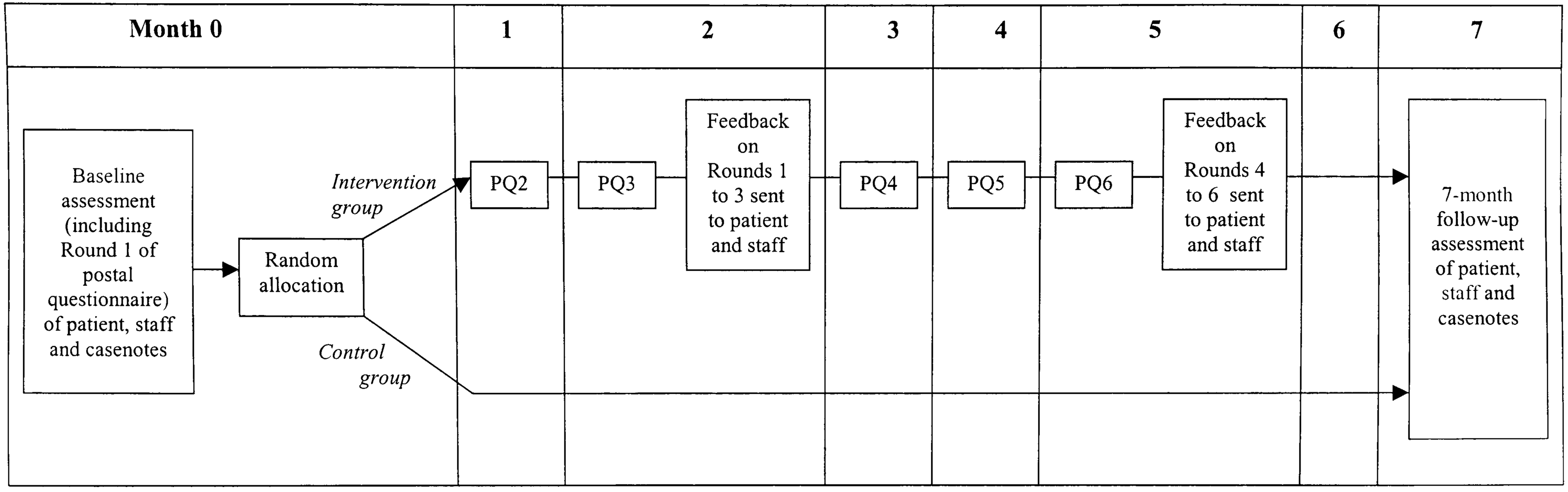
The patient follow-up assessment involved a second meeting with the researcher. The postal questionnaire was completed by the patient, and the time to complete the postal questionnaire along with the researcher's guess of allocation status were recorded. The evaluation measures were then administered.

The staff follow-up assessment involved either a face-to-face meeting with the researcher, or (especially where the staff had already done a follow-up assessment for another patient) leaving the follow-up form with the member of staff for completion and return by post. Administration time for the postal questionnaire element was recorded, either by the researcher (if present) or the member of staff. If the researcher was not present they made their allocation status guess after reading the postal questionnaire and before reading the evaluation measures.

An **Adverse Event form** was used to record any adverse events as they occurred during the study. Details of admissions were recorded on an **Admissions form** where they became known during the study. These data were augmented at the end of the study using the PAS and Current Clinical Summary (CCS) patient information systems. In practice, although the PAS community contact data were unreliable leading to sample selection problems (as described in Section 6.8), the PAS and CCS data were more reliable for admissions. The Admission form and the Adverse Event form are shown in Appendix 1.

A summary of the path through the study for an individual patient is shown in Figure 6.1.

Figure 6.1: Summary of intervention group and control group timings for individual patients



PQ2 = Round 2 postal questionnaire completed by patient and staff, PQ3 = Round 3 postal questionnaire completed by patient and staff, *etc*.

6.16 Protocol deviations

In addition to the protocol deviation in the sample selection method described in Section 6.8, one other change was required. The planned recruitment rate for the sample of 180 was 10 patients in month 1, then 30, 30, 30, 40 and 40 over the next 5 months. For a two-year study, this would allow the intervention to be provided for 18 months – six rounds of monthly questionnaires and three-monthly feedback. The recruitment rate achieved in practice for each month from May 2001 until November 2002 was 0, 7, 6, 5, 12, 18, 11, 14, 14, 6, 8, 16, 9, 10, 1, 4, 7, 7 and 5 – an average of 8.4 recruitments per month. The TSC were kept informed of the recruitment difficulties, and the following remedial actions were implemented:

1. Due to the better-than-expected retention rate, the target for recruitment was reduced from 180 to 160 (100 intervention, 60 control) with no loss of power (as noted in Section 6.5)
2. A request to pay staff £10 for each recruitment was made to the Local Research Ethics Committee in mid-2002. This request was denied.
3. The intervention was shortened to two rounds of monthly questionnaires and three-monthly feedback. Follow-up assessments were made at 7 months.

6.17 Analysis

Differences in administration time at baseline and follow-up were tested using paired sample t-tests. Differences between patients with and without follow-up primary outcome data were tested using chi-squared and independent samples t-test analysis. Since baseline scores were normally distributed, it was assumed that the same would be true of follow-up scores. The primary analytical strategy for testing the hypotheses was agreed in advance of viewing the follow-up data.

Hypothesis 1 (routine outcome assessment reduces unmet need) was tested using an independent samples t-test to compare patient-rated unmet need at follow-up for intervention and control group patients. As a sensitivity check (to consider changes due to time trends), the t-test was repeated with baseline data substituted for missing follow-up data. Hypothesis 2 (routine outcome assessment improves quality of life) was tested using an independent samples t-test to compare quality of life at follow-up for intervention and control group patients. As a sensitivity check (to consider changes due to time trends), the t-test was repeated with baseline data substituted for missing follow-

up data. Equal variances were assumed for all variables when comparing between control and intervention groups.

The investigation of interaction between premorbid IQ and responsiveness to the intervention involved univariate general linear modelling. The dependent variable was either patient-rated unmet need or quality of life. The fixed factors were allocation group (0=Control, 1=Intervention), gender (1=Male, 2=Female), CMHT (1 to 8), ethnicity (0=Non-White, 1=White), diagnosis (0=non-psychosis diagnosis, 1=psychosis diagnosis) and either the top half or top quarter premorbid IQ sub-group indicator (0=no, 1=yes). The covariates were age, baseline TAG, CANSAS-S met, CANSAS-S unmet, HAS-S, HAS-P, BPRS, HoNOS, CANSAS-P met and whichever of CANSAS-P unmet or MANSA was not the dependent variable. All variables were entered as main effects, and the 2-way interaction was tested between the NART sub-group indicator and allocation status.

Once evidence for the interaction was established, it was investigated further using linear regression. Three models were estimated, both with patient-rated unmet need and quality of life as the dependent variable. Due to the small number of cases, not all independent variables could be entered into the regression equation. Preliminary analysis indicated that age, gender and CMHT did not contribute to the model, so these were omitted. Model 1 included allocation status as the only independent variable. Model 2 entered variables in five blocks: ethnicity (White or Non-White), premorbid IQ (actual score), diagnosis (psychosis or not), all baseline clinical measures, and allocation status. Blocks 1 to 4 of Model 3 were the same as Model 2, block 5 comprised follow-up variables and block 6 comprised the allocation status. All variables in each block were entered together. Both r^2 and adjusted r^2 are reported.

The investigation of interaction between staff profession and responsiveness to the intervention also involved the fitting of interaction terms using univariate general linear modelling of follow-up patient-rated unmet need and quality of life. The fixed factors and covariates were the same as used to investigate premorbid IQ. All variables were entered as main effects, and the 2-way interaction was tested between the professional group and allocation status.

Hypothesis 3 (unmet need causes quality of life) was initially investigated by pooling data from the intervention and control groups. Analysis involved three stages:

1. Linear regression of follow-up quality of life with an independent variables of baseline patient-rated unmet need.
2. Linear regression of follow-up quality of life with independent variables of all sociodemographic and baseline clinical variables. Variables were entered in five blocks. Block 1 comprised sex (1=Male, 2=Female), age and ethnicity (0=Non-White, 1=White). Block 2 comprised premorbid IQ. Block 3 comprised seven dummy variables for the eight CMHTs. Block 4 comprised clinical diagnosis (0=non-psychosis diagnosis, 1=psychosis diagnosis). Block 5 comprised baseline assessments – CANSAS-P unmet, CANSAS-P met, MANSA, HAS-P, CANSAS-S unmet, CANSAS-S met, HAS-S, TAG, BPRS and HoNOS.
3. Linear regression of follow-up quality of life with independent variables of allocation status and all sociodemographic, baseline and follow-up clinical variables. Variables were entered in six blocks. The same first five blocks were used as previously, and block 6 comprised follow-up variables – CANSAS-P unmet, CANSAS-P met, CANSAS-S unmet, CANSAS-S met, HAS-P, HAS-S, TAG, BPRS and HoNOS.

This analysis was repeated with change scores, calculated by subtracting the baseline value from the follow-up value. Linear regression was repeated with change in MANSA as the dependent variable and the same first four blocks as previously. Block 5 comprised changes in CANSAS-P unmet, CANSAS-P met, CANSAS-S unmet, CANSAS-S met, HAS-P, HAS-S, TAG, BPRS and HoNOS. Percentages of variance are adjusted R^2 statistics, B is the regression coefficient, and Beta is the coefficient when the dependent and independent variables are standardised to have unit standard deviation.

Temporal precedence for changes in patient-rated unmet need and quality of life were investigated using cross-sectional time-series random-effects regression models for the intervention group only (since only this group had multivariate repeated measure data). The random-effects model takes a weighted average of the fixed-effects model (which allows for the fact that several ratings come from the same person) and the between-effects model (which models the mean rating from each person). Two models were

estimated. For the first model the dependent variable was MANSA score at month t, and the independent variables were CANSAS-P unmet need at months t and t-1. For the second model the dependent variable was change in MANSA score from month t-1 to month t, and the independent variables were change in CANSAS-P unmet need from month t-2 to t-1 and from month t-1 to month t. Percentages of variance are overall r^2 statistics.

Data analysis was undertaken according to an analysis plan established before viewing the data, using SPSS for Windows version 10 and Stata Version 8.

Chapter 7

Results

7.1 Baseline results

Recruitment began in May 2001, and ended in December 2002. 160 patients were recruited, comprising 101 in the intervention group and 59 in the control group. Patients from all eight Croydon CMHTs participated, ranging from 13 to 28 per CMHT. The characteristics of the patients are shown in Table 7.1.

Table 7.1: Clinical and social characteristics of patients (n=160)

	All (n=160)	**Intervention group (n=101)**	**Control group (n=59)**
Age (mean)	41.2 (s.d. 11.2)	41.8 (s.d. 11.4)	40.2 (s.d. 10.8)
Male	78 (49%)	48 (48%)	30 (51%)
Ethnicity			
White	122 (76%)	77 (76%)	45 (76%)
Black African-Caribbean	20 (13%)	16 (16%)	9 (15%)
Indian	6 (4%)	4 (4%)	2 (3%)
Other	12 (8%)	4 (4%)	3 (5%)
Highest educational level			
No formal qualification	61 (38%)	38 (38%)	23 (39%)
GCSE / GCE	45 (28%)	28 (28%)	19 (32%)
A levels or equivalent	14 (9%)	10 (10%)	3 (5%)
Higher diploma or degree	16 (10%)	11 (11%)	4 (7%)
Not known	24 (15%)	13 (13%)	10 17%)
Primary clinical diagnosis			
Schizophrenia	60 (38%)	40 (40%)	20 (34%)
Bipolar affective disorder	17 (11%)	8 (8%)	9(15%)
Other psychoses	21 (13%)	12 (12%)	7 (12%)
Affective disorder	43 (27%)	27 (27%)	16 (27%)
Personality disorder	11 (7%)	7 (7%)	4 (7%)
Other	8 (5%)	7 (7%)	3 (5%)
Contact with mental health services			
Years since first contact	13.1 (s.d. 11.8)	14.2 (s.d. 12.6)	11.1 (s.d. 9.8)
Years in this episode of care	4.1 (s.d. 4.2)	4.3 (s.d. 4.7)	3.9 (s.d. 3.3)

74 staff participated in baseline assessments, comprising 43 psychiatric nurses, 14 care managers, 11 psychiatrists, 3 occupational therapists, 1 clinical psychologist, 1 counselling psychologist, and 1 Independent Living Officer. In total, 102 baseline staff assessments were completed by psychiatric nurses, 28 by care managers, 21 by psychiatrists, 3 by psychologists, 4 by occupational therapists and 2 by an Independent Living Officer.

The postal questionnaire element of the baseline assessment took a mean of 14.9 minutes for patients and 7.8 minutes for staff. The baseline scores for the assessments are shown in Table 7.2.

Table 7.2: Baseline measures for the FOCUS RCT (n=160)

Measure	**All (n=160)** mean (s.d.)	**Intervention (n=101)** mean (s.d.)	**Control (n=59)** mean (s.d.)
Staff-completed			
CANSAS-S unmet	2.98 (3.19)	3.24 (3.31)	2.54 (2.94)
CANSAS-S met	5.04 (3.43)	5.06 (3.29)	5.02 (3.69)
TAG	5.21 (3.64)	5.44 (3.58)	4.81 (3.73)
HAS-S	7.34 (1.61)	7.45 (1.59)	7.14 (1.64)
HoNOS	8.87 (6.43)	9.15 (6.63)	8.40 (6.10)
Patient-rated			
CANSAS-P unmet	4.59 (3.62)	4.36 (3.36)	4.98 (4.05)
CANSAS-P met	4.21 (2.88)	4.23 (2.81)	4.17 (3.04)
HAS-P	7.95 (1.94)	8.19 (1.79)	7.54 (2.12)
MANSA	4.25 (1.01)	4.25 (0.99)	4.25 (1.05)
Interviewer-rated			
BPRS	33.51 (9.29)	33.35 (9.04)	33.79 (9.78)
NART error score	25.4 (11.90)	25.77 (12.22)	24.88 (11.47)
Converted into premorbid IQ	99	99	100

In addition, the baseline Supplementary-P form indicated 124 (78%) had a close friend (comprising 81 (80%) in the intervention group and 43 (73%) in the control group), 105 (66%) had seen a friend in the last week (comprising 74 (73%) in the intervention group

and 31 (53%) in the control group), 16 (10%) had been accused of a crime (comprising 8 (8%) in the intervention group and 8 (14%) in the control group) and 26 (17%) had been a victim of physical violence (comprising 17 (17%) in the intervention group and 9 (15%) in the control group). No comparison was made between any intervention and control group baseline assessments, since any differences would have arisen by chance (Everitt & Wessely, 2004).

7.2 Intervention provision

The intervention was offered to all 101 patients in the intervention group. Details of the number of questionnaires sent out and returned are shown in Table 7.3.

Table 7.3: Postal questionnaire distribution and return rates

	Round 2	**Round 3**	**Round 4**	**Round 5**	**Round 6**	**Mean**
Staff sent (n)	100	98	97	96	95	97.2
Staff returned (n)	78	70	65	57	55	65.0
Staff not returned (n)	22	28	32	39	40	32.2
Staff response rate (%)	**78**	**71**	**67**	**59**	**58**	
Patient sent (n)	99	99	98	96	95	97.4
Patient returned (n)	84	83	74	73	72	77.2
Patient not returned (n)	15	16	24	23	23	20.2
Patient response rate (%)	**85**	**84**	**76**	**76**	**76**	

As requested, some staff returned blank postal questionnaires when they had not seen the patient since the last round of assessment and had received no new information about them. A total of 25 blank questionnaires were received during the study, comprising six in round 2, four in round 3, six in round 4, five in round 5, and four in round 6. For each blank questionnaire, the staff ratings from the previous round were used in the analysis. This is because blank questionnaires did not constitute missing data (which would have been left blank), but rather provided information that the staff assessment was unchanged from previously.

The sending of postal questionnaires was discontinued when a patient died (n=2, both from natural causes), the patient withdrew consent for further involvement (n=1), or no staff could be identified as in contact with the patient (n=3). Due to this attrition, a total

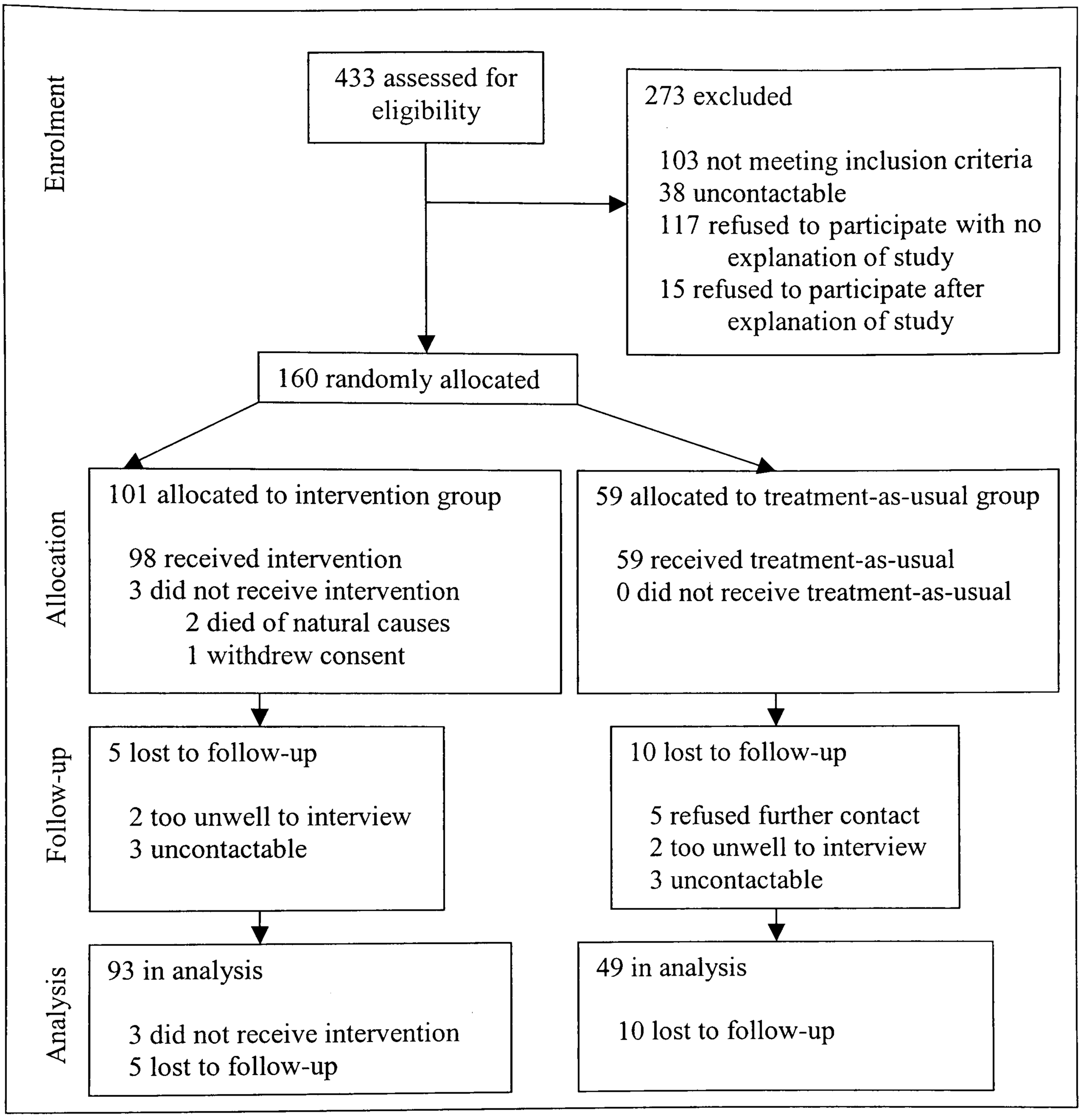
of 486 (96%) out of a planned 505 staff postal questionnaires were sent out. Of these 486, 325 were returned, giving a staff response rate of 67%. 487 postal questionnaires were sent out to patients, and 386 returned, giving a patient response rate of 79%.

Three-monthly summary feedback was sent after Round 3 to 96 staff-patient pairs (95% of intervention group), and after Round 6 to 93 staff-patient pairs (92% of intervention group). Since the assessment and the feedback are the active ingredients of the intervention according to the FOCUS Model, this indicates that 92% of the intervention group received the full intervention, and the remaining 8% received only a partial intervention.

7.3 Retention

Follow-up assessments were completed by July 2003. Follow-up data were collected for 93 (92%) patients in the intervention group and 49 (83%) patients in the control group, a total of 142 (89%) out of 160 patients in the study. The trial flow diagram is shown in Figure 7.1, using the Consolidated Standards of Reporting Trials (CONSORT) format (Altman et al, 2001).

Figure 7.1: Trial flow diagram for patients in the FOCUS RCT



A total of 18 patients did not have follow-up ratings for the primary outcome measure of patient-rated unmet need, and 19 for quality of life (one patient completed CANSAS but not MANSA). Of the 18 missing both, 6 were male, 13 were white, 13 had a psychotic diagnosis, mean age was 41.3 (s.d. 11.2) and mean premorbid IQ was 98.3 (s.d. 13.7). Mean staff-rated baseline score for unmet need was 2.8 (s.d. 2.5), for met need was 4.7 (s.d. 3.9), for HAS-S was 6.8 (s.d. 1.2), for TAG was 5.3 (s.d. 3.1), and for HoNOS 8.8 (s.d. 5.0). Mean patient-rated baseline score for unmet need was 3.2 (s.d. 3.4), for met need was 3.3 (s.d. 2.1), for MANSA was 4.6 (s.d. 1.3), for HAS-P was 7.9 (s.d. 1.8), and for BPRS was 33.1 (s.d. 13.1). None of the sociodemographic or baseline

clinical variables differed between the 142 patients with and the 18 patients without full follow-up data.

Two intervention group patients were only found to be too unwell to interview when a face-to-face interview was being requested by the researcher. These patients were still sent the postal questionnaires and feedback (and hence deemed to have received the intervention) even though they may have been too unwell to make use of it. The intervention-group patient who withdrew consent during the study stated that the questions were "too disturbing and intrusive". This withdrawal of consent was notified to the Trial Steering Committee as an adverse event.

One intervention-group patient was sent to prison on remand during the intervention, following a serious assault. Although the intervention was discontinued, they were interviewed for follow-up (prematurely, in the context of contact by his clinical team), and since complete follow-up data were available, this patient was included in the analysis. There was no evidence linking the assault with involvement in the study, and the Trial Steering Committee was informed of this adverse event.

Follow-up staff data were collected for 56 (95%) control and 95 (94%) intervention group patients, a total of 151 (94%) out of 160 patients in the study. Reasons for non-collection for the three control group patients were that no staff could be identified as in contact (n=2) and transfer to a new member of staff who refused due to already having several patients in the study (n=1). Reasons for non-collection for the six intervention group patients were that no staff were identified as in contact (n=2), transfer to a new member of staff who refused due to already having several patients in the study (n=2), staff uncontactable (n=1) and the patient died (n=1). By the end of their time in the study, 41 (26%) patients had a different staff member than at baseline.

Follow-up casenote data were collected for 58 (98%) control and 100 (99%) intervention group patients, a total of 158 (99%) out of 160 patients in the study. Both missing casenotes were due to administrative errors by the research team.

7.4 Follow-up results

The follow-up scores for each measure are shown in Table 7.4. Two patients (1 intervention, 1 control) were in hospital when they started in the study.

Table 7.4: Follow-up measures for the FOCUS RCT

Measure	All mean (s.d.)	Intervention group mean (s.d.)	Control group mean (s.d.)	Difference	95% CI
Staff-completed	*n=151*	*n=95*	*n=56*		
CANSAS-S unmet	2.66 (3.25)	2.84 (3.49)	2.34 (2.79)	-0.50	-1.6 to 0.6
CANSAS-S met	4.56 (3.29)	4.26 (3.03)	5.07 (3.67)	0.81	-0.3 to 1.9
TAG	5.16 (3.51)	5.14 (3.51)	5.20 (3.54)	0.06	-1.1 to 1.2
HAS-S	7.35 (1.72)	7.51 (1.57)	7.08 (1.92)	-0.43	-1.0 to 0.1
HoNOS	9.43 (6.42)	9.35 (6.45)	9.55 (6.42)	0.19	-2.0 to 2.4
Patient-rated	*n=142*	*n=93*	*n=49*		
CANSAS-P unmet	4.01 (3.83)	3.96 (3.58)	4.10 (4.31)	0.15	-1.2 to 1.5
CANSAS-P met	4.47 (3.84)	4.39 (3.32)	4.63 (4.71)	0.25	-1.1 to 1.6
HAS-P	7.28 (2.23)	7.37 (2.15)	7.12 (2.38)	-0.25	-1.0 to 0.5
MANSA	4.24 (1.07)	4.27 (1.04)	4.20 (1.14)	-0.07	-0.4 to 0.3
Interviewer-rated	*n=142*	*n=93*	*n=49*		
BPRS	31.85 (10.03)	31.39 (9.27)	32.71 (11.39)	1.3	-2.2 to 4.8
Service use	*n=160*	*n=101*	*n=59*		
Hospital bed days	9.86 (30.61)	7.74 (25.35)	13.49 (37.93)	5.75	-4.2 to 15.7

No intervention and control group scores differed significantly. There was also no significant difference when hospital admission status (having versus not having any hospital bed days during the study period) was compared between the groups (Chi-squared<1, df=1, p=ns).

In addition, the follow-up Supplementary-P form indicated 112 (79%) had a close friend (comprising 73 (79%) in the intervention group and 39 (80%) in the control group), 92 (65%) had seen a friend in the last week (comprising 64 (69%) in the intervention group and 28 (57%) in the control group), 10 (7%) had been accused of a crime (comprising 7 (8%) in the intervention group and 3 (6%) in the control group) and 19 (14%) had been a victim of physical violence (comprising 17 (19%) in the intervention group and 2 (4%) in the control group). Since the time frame for some of these items was one year (*i.e.* longer than the time in the study), no comparison was made between allocation groups.

The postal questionnaire element of the follow-up assessments took a mean of 8.7 minutes for patients and 7.4 minutes for staff. There was a significant reduction in completion time by the 129 patients for whom baseline and follow-up completion time data were available ($t=8.5$, $df=128$, $p<0.001$), but not for the 130 staff with these data.

After the postal questionnaire element of the follow-up assessment, the researcher recorded whether they believed the interviewee (staff or patient) was in the intervention group, control group or did not know. This rating was recorded for 140 patient and 143 staff interviews, and the results are shown in Table 7.5.

Table 7.5: Researcher rating of allocation (n=143)

	Researcher rating of staff allocation				*Researcher rating of patient allocation*			
	n	**Intervention**	**Control**	**Not known**	**n**	**Intervention**	**Control**	**Not known**
True allocation								
Intervention	89	36	2	51	92	61	1	30
Control	54	5	19	30	48	3	34	11
Total	143	41	21	81	140	64	35	41

Some researcher blinding to allocation status was retained. In 81 (57%) of the 143 staff interviews and in 41 (29%) of the 140 patient interviews the researcher was unable to guess allocation status. Where they did rate allocation status, they were correct for 97 (92%) of their 105 intervention group ratings, and for 53 (95%) of their 56 control group ratings.

7.5 Testing of hypothesis 1

Hypothesis 1 proposed that receiving the intervention would be associated with a decrease of at least 1.0 patient-rated unmet needs, as measured using CANSAS-P. For the 142 patients with baseline and follow-up patient-rated unmet need data, 79 (56%) had at least 1 fewer unmet need at follow-up, comprising 51 (55%) out of 93 in the intervention group and 28 (57%) out of 49 in the control group.

An independent samples t-test on follow-up patient-rated unmet need (based on the assumption of no baseline differences in patient-rated unmet need between groups) indicated no difference between groups in follow-up patient-rated unmet need (mean difference = 0.15, t=0.21, df=140, p=0.83). As a sensitivity check, baseline values were substituted for missing follow-up patient-rated unmet need cases. The t-test again indicated no differences between groups (mean difference = 0.04, t=0.09, df=157, p=0.93). To increase precision in the estimate of responsiveness to the intervention, a regression of follow-up patient-rated unmet need with allocation status and baseline patient-rated unmet need (*i.e.* ANCOVA) was undertaken. This showed a strong association between baseline and follow-up patient-rated unmet need (B=0.56, Beta=0.53, p<0.001), but no association with allocation status (B=0.34, Beta=0.04, p=0.56).

The odds of significant improvement in the intervention group is (51/42) = 1.21, and the odds in the control group is (28/21) = 1.33. The odds ratio is therefore (1.21/1.33) = 0.91 (95% CI 0.45 to 1.83), indicating no significant effect size. No association between allocation status and follow-up patient-rated unmet need was found. The null hypothesis, that there is no difference between the intervention and control groups, cannot be rejected.

7.6 Testing of hypothesis 2

Hypothesis 2 proposed that receiving the intervention would be associated with an increase of at least 0.25 points in quality of life, as measured using MANSA. For the 141 patients with baseline and follow-up quality of life data, 56 (40%) had a MANSA rating at least 0.25 higher at follow-up, comprising 39 (42%) out of 92 in the intervention group and 17 (35%) out of 49 in the control group.

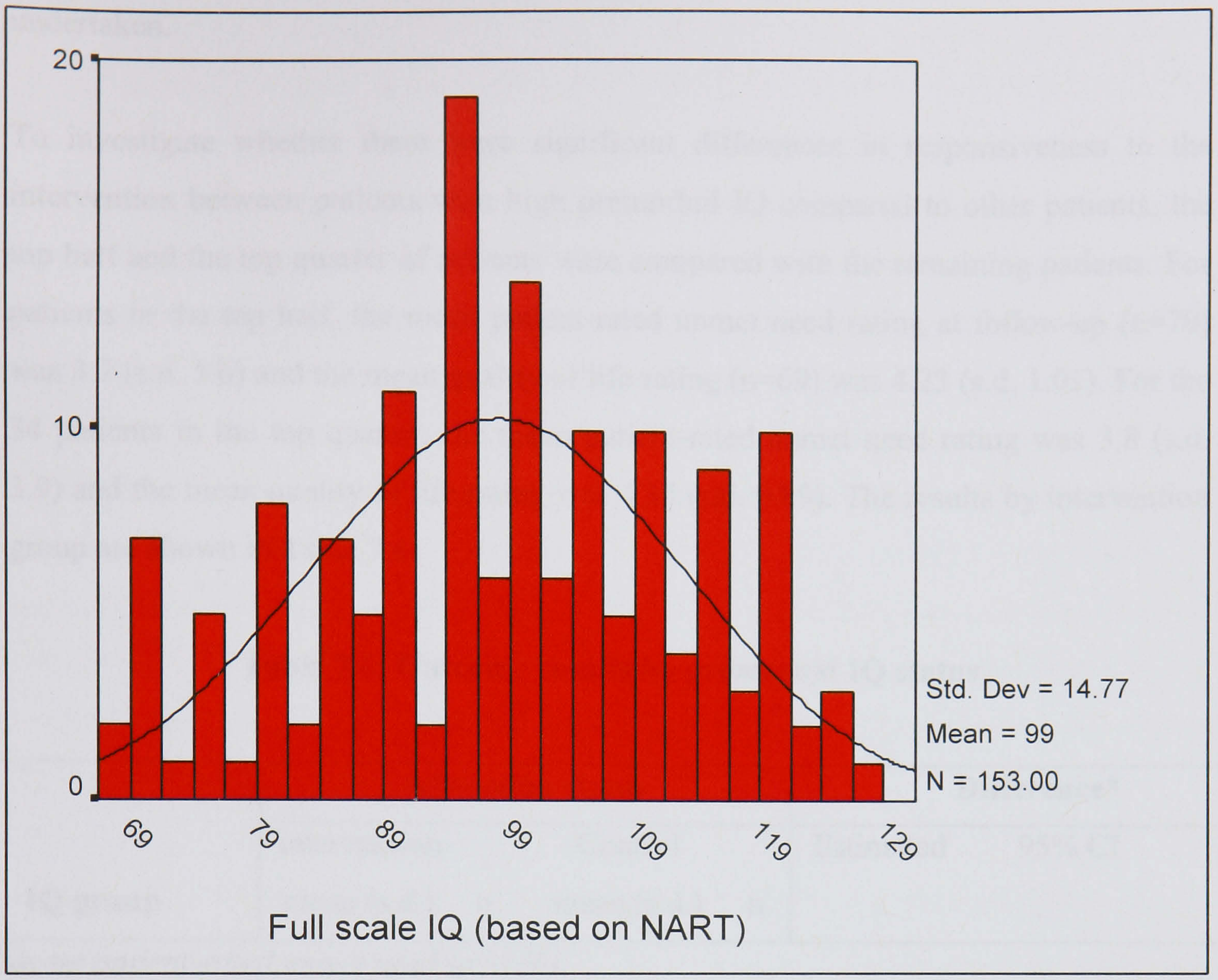
An independent samples t-test on follow-up quality of life (based on the assumption of no baseline differences in quality of life between groups) indicated no difference between groups in follow-up quality of life (mean difference = 0.07, t=0.36, df=139, p=0.72). As a sensitivity check, baseline values were substituted for missing follow-up quality of life cases, to allow intention-to-treat analysis. The t-test again indicated no differences between groups (mean difference = 0.05, t=0.28, df=158, p=0.78). Finally, to increase precision in the estimate of responsiveness to the intervention, a regression of follow-up quality of life with allocation status and baseline quality of life was undertaken. This showed a strong association between baseline and follow-up quality of life (B=0.76, Beta=0.69, p<0.001), but no association with allocation status (B=-0.02, Beta=-0.01, p=0.87).

The odds of significant improvement in the intervention group is (39/53) = 0.74, and the odds in the control group is (17/32) = 0.53. The odds ratio is therefore (0.74/0.53) = 1.39 (95% CI 0.67 to 2.84), indicating no significant effect size. No association between allocation status and follow-up quality of life was found. The null hypothesis, that there is no difference between the intervention and control groups, cannot be rejected.

7.7 *A priori* sub-group analysis 1: Premorbid IQ

The premorbid IQ scores for the 153 patients with completed NART data ranged from 69 to 128. The distribution of scores is shown in Figure 7.2.

Figure 7.2: Distribution of premorbid Full-scale IQ scores



The 153 patients were partitioned into quartiles, comprising those with premorbid IQ scores ranging from 69 to 89 (n=38), from 90 to 99 (n=38), from 100 to 110 (n=39), and from 111 to 128 (n=38). They were also partitioned into halves, comprising those ranging from 69 to 99 (n=76) and from 100 to 128 (n=77). The patient sub-groups comprising the top quartile (n=38) and the top half (n=77) were investigated. Of the 153 patients with premorbid IQ scores, 135 (88%) also had baseline and follow-up quality of life, and 136 (89%) also had baseline and follow-up patient-rated unmet need.

To test for an interaction between allocation status and premorbid IQ scores, models including interaction terms were fitted. The interaction terms were between allocation status and either being in the top half or in the top quartile patient sub-groups. The dependent variable was either follow-up patient-rated unmet need or quality of life. Four models were therefore estimated. Each model used all sociodemographic and baseline clinical variables. An interaction was found for follow-up patient-rated unmet need between allocation status and the top quarter of patients (p=0.002). An interaction was found for quality of life between allocation status and the top half of patients (p=0.015). Since this analysis provided some *post hoc* evidence that the intervention

may be differentially effective on the basis of premorbid IQ, further analysis was undertaken.

To investigate whether there were significant differences in responsiveness to the intervention between patients with high premorbid IQ compared to other patients, the top half and the top quarter of patients were compared with the remaining patients. For patients in the top half, the mean patient-rated unmet need rating at follow-up (n=70) was 3.7 (s.d. 3.6) and the mean quality of life rating (n=69) was 4.23 (s.d. 1.01). For the 34 patients in the top quarter, the mean patient-rated unmet need rating was 3.8 (s.d. 3.9) and the mean quality of life rating was 4.34 (s.d. 0.89). The results by intervention group are shown in Table 7.6.

Table 7.6: Outcome results by premorbid IQ status

	Allocation status				**Difference***		
IQ group	Intervention mean (s.d.)	n	Control mean (s.d.)	n	Estimated	95% CI	p
Follow-up patient-rated unmet need (n=136)							
Top half	3.3 (3.3)	47	4.7 (4.1)	23	1.4	-0.4 to 3.2	0.13
Lower half	4.6 (3.8)	43	4.0 (4.7)	23	-0.7	-2.8 to 1.5	0.55
Top quarter	2.5 (2.6)	21	5.9 (4.7)	13	**3.4**	**0.8 to 5.9**	**0.012**
Lower three quarters	4.4 (3.7)	69	3.7 (4.1)	33	-0.7	-2.3 to 1.0	0.43
Follow-up quality of life (n=135)							
Top half	4.4 (1.0)	46	3.8 (1.0)	23	**-0.6**	**-1.1 to –0.1**	**0.02**
Lower half	4.1 (1.1)	43	4.4 (1.1)	23	0.3	-0.3 to 0.8	0.37
Top quarter	4.5 (0.7)	21	4.0 (1.1)	13	-0.5	-1.2 to 0.1	0.09
Lower three quarters	4.2 (1.1)	68	4.2 (1.1)	33	0.0	-0.5 to 0.4	0.83

* Difference = control score - intervention score

There were significant differences between the sub-groups in terms of their responsiveness to the intervention – quality of life appears to be improved by the intervention for the top half of patients, and patient-rated unmet needs appear to be reduced by the intervention for the top quarter.

For patients in the top quarter of premorbid IQ, 13 (62%) of the 21 intervention group reduced their patient-rated unmet need by at least 1.0 needs, as did 4 (31%) of the 13 control group. This gives an odds ratio of ((13/8)/(4/9)) = 3.66 (95% CI 0.84 to 15.91). The Number Needed to Treat (NNT) is 1/(Proportion of events in intervention group - proportion of events in control group) = 1/((13/21)-(4/13)) = 3, indicating (if the effect size is replicated in a more powered study) 3 patients with a premorbid IQ of at least 111 would need to receive the intervention for one to experience a reduction of 1.0 patient-rated unmet needs.

For patients in the top half of premorbid IQ, 25 (54%) of the 46 intervention group improved in MANSA rating of quality of life by at least 0.25, as did 7 (30%) of the 23 control group. This gives an odds ratio of ((25/21)/(7/16)) = 2.72 (95% CI 0.94 to 7.86), and a NNT of 1/((25/46)-(7/23)) = 4, indicating 4 patients with a premorbid IQ of at least 100 would need to receive the intervention for one to experience an increase of 0.25 on MANSA rating of quality of life.

The impact of including baseline and follow-up variables was investigated using linear regression. Three linear regression models were estimated. Model 1 included allocation status as the only independent variable, to confirm that there was a difference in responsiveness to the intervention between patient sub-groups. Model 2 included allocation status, sociodemographic and baseline clinical measures. Model 3 included all independent variables from Model 2 plus follow-up variables.

For Model 1, the results confirmed those shown in Table 7.6. Allocation status was a predictor of patient-rated unmet need for the top quarter (B=-3.4, p=0.01) but not the lower three quarters (B=0.7, p=0.43). Allocation status was not a predictor of quality of life for the top quarter (B=0.5, p=0.09) or the lower three quarters (B=0.0, p=0.83). Allocation status was a predictor of quality of life for the top half (B=0.6, p=0.02) but not the lower half (B=-0.3, p=0.37). Allocation status was not a predictor of patient-rated unmet need for the top half (B=-1.4, p=0.13) or the lower half (B=0.6, p=0.54).

For Models 2 and 3, non-significant results in Table 7.6 were all re-confirmed. The results for the lower premorbid IQ patient sub-groups did not indicate any response to the intervention. Allocation status did not predict follow-up quality of life for the lower half (Model 2: B=-0.5, p=0.06; Model 3: B=-0.1, p=0.73), or for the lower three

quarters (Model 2: B=-0.1, p=0.59; Model 3: B=0.1, p=0.74). Allocation status also did not predict follow-up patient-rated unmet need for the lower half (Model 2: B=2.1, p=0.09; Model 3: B=0.7, p=0.56) or the lower three quarters (Model 2: B=1.4, p=0.07; Model 3: B=1.1, p=0.13). Finally, allocation status did not predict follow-up patient-rated unmet need for the top half (Model 2: B=-0.6, p=0.48; Model 3: B=0.0, p=0.96), or follow-up quality of life for the top quarter (Model 2: B=0.4, p=0.11; Model 3: B=-0.7, p=0.18).

One positive result from Table 7.6 was not supported. Allocation status did not predict follow-up quality of life for the top half (Model 2: B=0.3, p=0.10; Model 3: B=0.2, p=0.21), indicating that the relationship found in Table 7.6 is not maintained when controlling for the influence of baseline or follow-up variables. The other positive result – allocation status predicts follow-up patient-rated unmet need for the top quarter – was confirmed using linear regression, shown in Table 7.7.

Table 7.7: Regression on follow-up patient-rated unmet need for premorbid IQ top quarter of patients (n=34), with (Model 3) and without (Model 2) follow-up variables

	Model 2 (baseline only)			Model 3 (baseline and follow-up)		
	B	**Beta**	**p**	**B**	**Beta**	**p**
Allocation status	**-3.33**	**-0.41**	**0.004**	**-3.88**	**-0.48**	**0.047**
Ethnicity	**3.53**	**0.33**	**0.020**	2.01	0.19	0.16
Premorbid IQ	0.00	0.00	0.97	-0.18	-0.22	0.13
Psychosis	1.96	0.24	0.10	**3.20**	**0.39**	**0.026**
Baseline						
CANSAS-P unmet	**0.72**	**0.70**	**0.002**	**0.62**	**0.59**	**0.008**
CANSAS-P met	0.09	0.08	0.58	-0.13	-0.11	0.37
MANSA	1.40	0.34	0.12	**3.73**	**0.88**	**0.005**
HAS-P	-0.47	-0.22	0.14	-0.72	-0.34	0.09
CANSAS-S unmet	-0.29	-0.18	0.41	-0.50	-0.31	0.32
CANSAS-S met	-0.31	-0.25	0.09	-0.28	-0.22	0.18
HAS-S	**0.93**	**0.43**	**0.026**	-0.01	-0.01	0.97
TAG	0.51	0.35	0.18	-0.56	-0.39	0.24
BPRS	0.12	0.26	0.22	0.00	0.01	0.97
HoNOS	0.06	0.07	0.83	**0.78**	**0.97**	**0.046**
Follow-up						
MANSA				**-3.32**	**-0.76**	**0.011**
CANSAS-P met				**0.70**	**0.57**	**0.018**
HAS-P				-0.52	-0.28	0.09
CANSAS-S unmet				-0.84	-0.41	0.06
CANSAS-S met				0.01	0.01	0.94
HAS-S				**1.38**	**0.59**	**0.049**
TAG				-0.06	-0.04	0.81
BPRS				-0.13	-0.35	0.08
HoNOS				**0.41**	**0.55**	**0.021**

Bold = $p<0.05$

Receiving the intervention was associated with reduced patient-rated unmet need even when controlling for baseline and follow-up variables. For Model 2, r^2 was 0.79 and adjusted-r^2 was 0.63. For Model 3, r^2 was 0.97 and adjusted-r^2 was 0.89. The effect size is given by ((intervention group mean - control group mean) / standard deviation of both groups) = (2.5-5.9)/3.9 = -0.8, a moderate to large effect size.

7.8 *A priori* sub-group analysis 2: Staff group

Three staff groups had more than 5 members in both the intervention and control groups: psychiatrists (11 intervention, 6 control, 17 in total), psychiatric nurses (33 intervention, 60 control, 93 in total) and care managers (17 intervention, 7 control, 24 in total). To test for an interaction between allocation status and staff profession, models including interaction terms were fitted. An interaction was found for follow-up patient-rated unmet need between allocation status and professional group (p=0.028). No interaction was found for quality of life between allocation status and professional group (p=0.33). This analysis provided some evidence that the intervention may be differentially effective on the basis of staff group, so further analysis was undertaken.

To investigate whether there was any difference between the professions in the extent to which they 'made use of' the intervention, change scores (follow-up score – baseline score) for the two primary outcomes were calculated. The mean change in patient-rated unmet need for the three staff groups is shown in Table 7.8.

Table 7.8: Change in follow-up patient-rated unmet need, by profession

	Intervention*	**Control ***	**Difference**		
	mean (s.d.)	mean (s.d.)	mean	95% CI	p
Psychiatric nurses	-1.17 (2.54)	-0.85 (4.79)	0.32	-1.19 to 1.82	0.68
Care managers	0.47 (2.81)	-1.00 (1.91)	-1.47	-3.89 to 0.95	0.22
Psychiatrists	1.45 (4.99)	-3.33 (2.94)	**-4.79**	**-9.56 to -0.01**	**0.049**

* Change = Follow-up score - baseline score, so negative score indicates improvement

The mean change in quality of life for each professional group is shown in Table 7.9.

Table 7.9: Change in follow-up quality of life, by profession

	Intervention*	**Control ***	**Difference**		
	mean (s.d.)	mean (s.d.)	mean	95% CI	p
Psychiatric nurses	0.21 (0.59)	0.02 (0.80)	-0.19	-0.48 to 0.10	0.20
Care managers	-0.45 (1.02)	0.12 (1.31)	0.58	-0.45 to 1.60	0.26
Psychiatrists	-0.26 (0.91)	0.14 (1.14)	0.40	-0.67 to 1.47	0.44

* Change = Follow-up score - baseline score, so positive score indicates improvement

Given the turnover (26%) of staff during the intervention and the small numbers for each profession, no further analysis was undertaken by profession.

7.9 Evaluation of the FOCUS Model

The FOCUS Model was investigated at follow-up using the Impact of Involvement questionnaire, described in Section 6.10.2. It comprises a self-rated assessment of responses to the questions shown in Table 7.10, with separate versions for staff and patients. The measure was developed to investigate whether the cognitive and behavioural sequelae predicted by the FOCUS Model did, in fact, occur. The purpose was to validate the FOCUS Model.

The Impact of Involvement questionnaire was completed by 122 staff (81 intervention, 41 control) and 125 patients (85 intervention, 40 control). Control group participants were also asked to complete the questionnaire (since the researcher was blind to allocation status), but only responses from the 81 (80%) staff and 85 (84%) patients in the intervention group were considered. A summary of their ratings is shown in Table 7.10. Some questions were left blank, so each percentage shown is derived from the total number of valid answers for the question. Parts of the question shown in bold in Table 7.10 were emboldened in the questionnaire, to emphasise what was being asked.

Table 7.10: Intervention group staff (n=81) and patient (n=85) responses to Impact of Involvement questionnaire

Question*	Number (%) replying 'Yes'	
	Staff	Patient
Did **filling in the FOCUS questionnaires** make you think about the care *the service user gets*?	72 (94)	69 (81)
Did **filling in the FOCUS questionnaires** make you think about your relationship with the *service user*?	71 (92)	60 (71)
Did you receive the feedback?	70 (88)	80 (94)
Did you read the feedback?	69 (96)	70 (85)
Did you understand the feedback?	61 (88)	69 (84)
Did **receiving the feedback** make you think about the care *the service user is receiving*?	59 (82)	52 (64)
Did **receiving the feedback** make you think about your relationship with *the service user*?	60 (85)	53 (65)
Did receiving the feedback lead you to discuss the content of *their* care with *the service user*?	36 (51)	26 (31)
Did receiving the feedback lead you to change your behaviour with *the service user*?	30 (41)	13 (16)

* In the patient version, italicised wording was altered to refer to staff

Combining the final two questions, there was self-reported evidence of behaviour change from 46 (63%) staff and 29 (36%) patients.

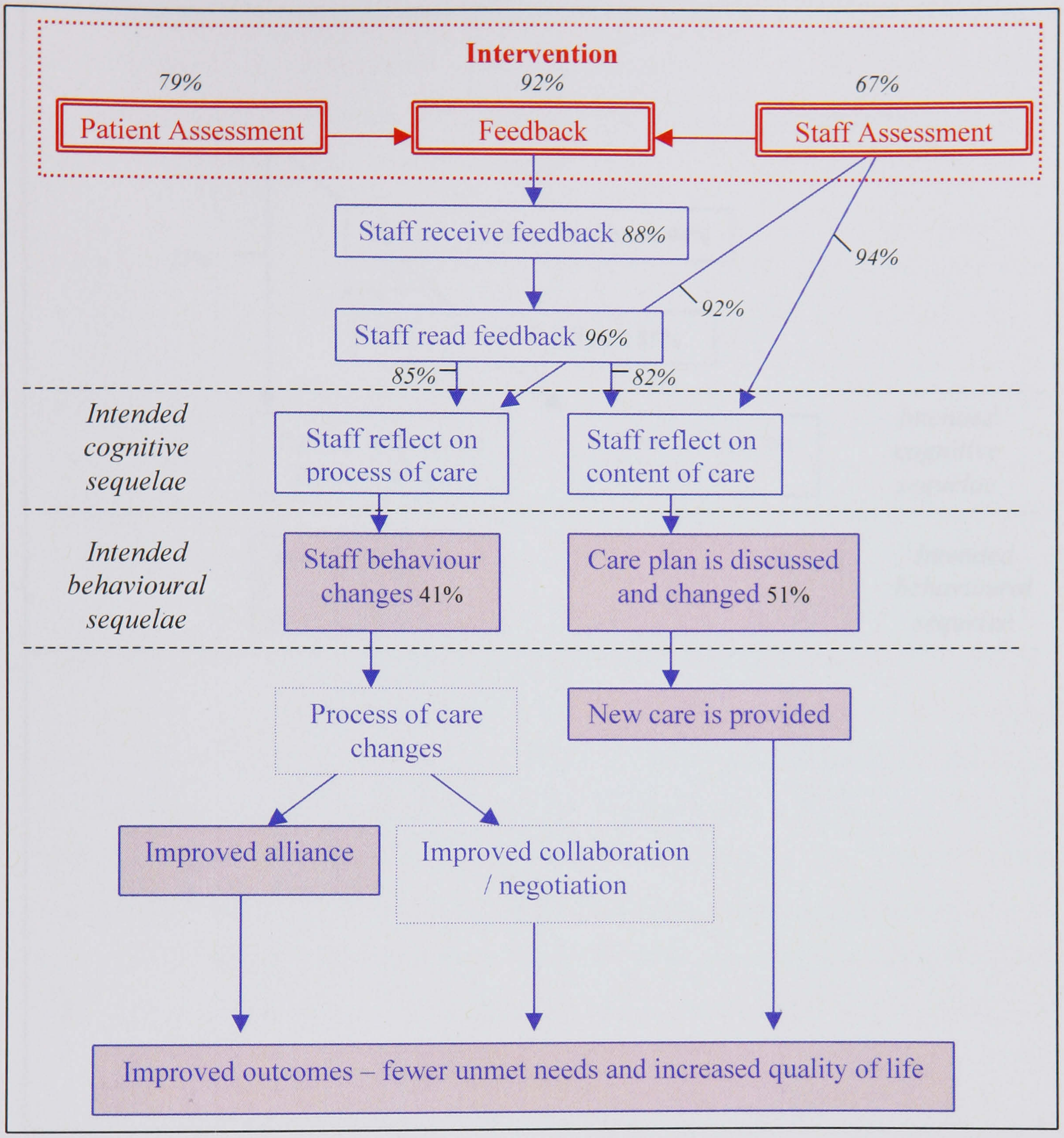
The care plan was rated at baseline and follow-up for pre-specified components of direct care, planned assessment, planned liaison and carer support. The results are shown in Table 7.11.

PAGE NUMBERING AS ORIGINAL

which were not validated are indicated by greyed boxes, and components which were not directly tested are indicated by dotted-line boxes.

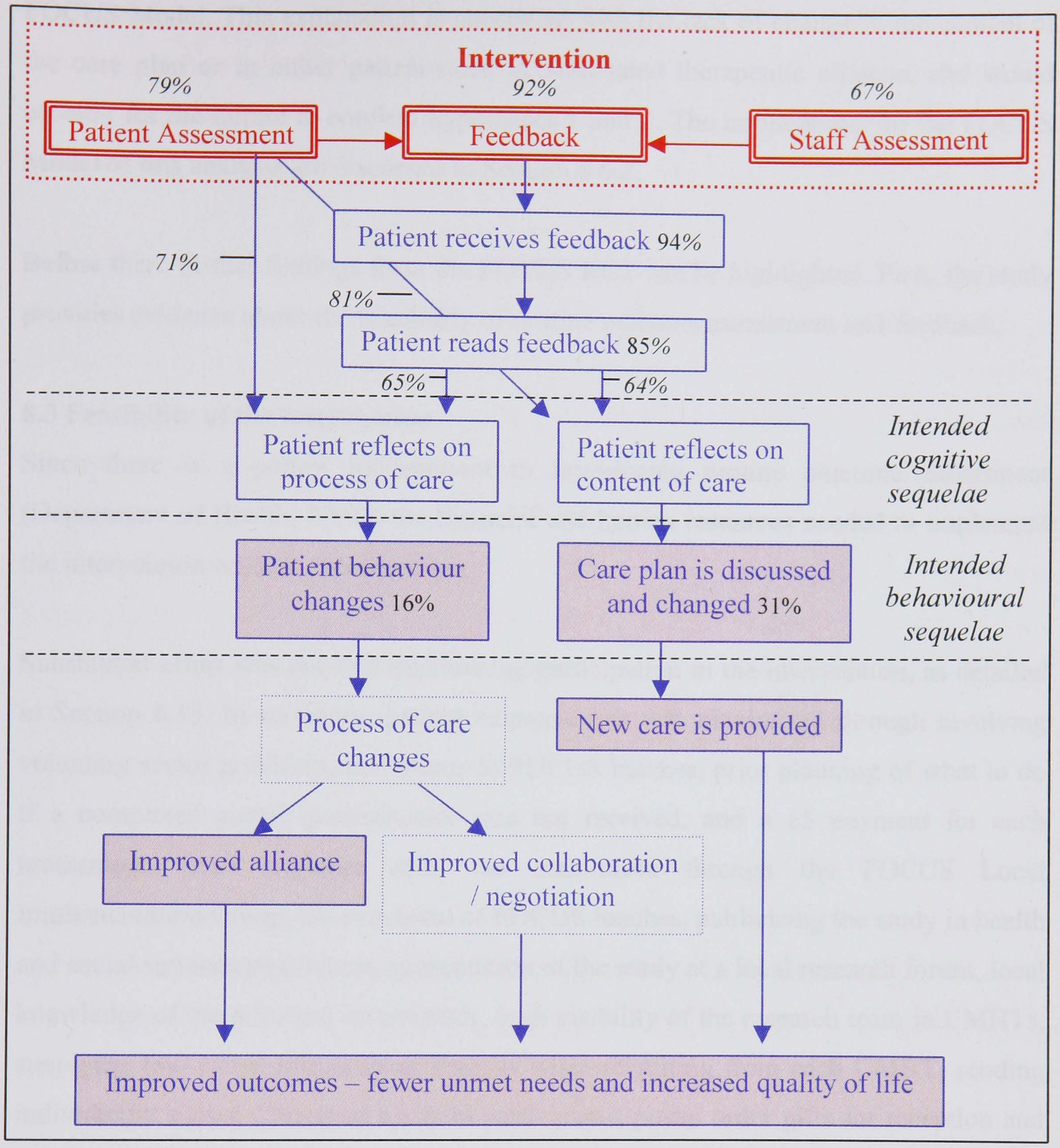
Evidence for the staff version of the FOCUS Model (originally depicted in Figure 4.1) is shown in Figure 8.1, with annotation and dashed lines used to indicate more clearly the intended cognitive and behavioural sequelae of receiving the intervention.

Figure 8.1: Evidence from the FOCUS RCT for the staff version of the FOCUS Model



A summary of the evidence for the patient version of the FOCUS Model (as depicted in Figure 4.2) is shown in Figure 8.2.

Figure 8.2: Evidence from the FOCUS RCT for the patient version of the FOCUS Model



Figures 8.1 and 8.2 suggest that the predicted cognitive sequelae of the intervention did take place, but the predicted behavioural sequelae did not. Specifically, 37% of staff and 64% of patients reported no behavioural sequelae. Given that acquiescence (social desirability) bias is likely to have inflated the number of positive responses retrospectively given by the interviewee about the impact of the study, this self-report approach may over-estimate the impact of participation. Therefore the true level of behaviour change may be even lower.

Overall, the weight of evidence favours the conclusion that the intervention was insufficiently powerful to cause the intended behavioural sequelae described in the FOCUS Model. This explanation is consistent with the lack of change in the content of the care plan or in either patient-rated or staff-rated therapeutic alliance, and would account for the failure to confirm hypotheses 1 and 2. The implications for the FOCUS Model of this analysis are discussed in Section 8.6.2.

Before then, further findings from the FOCUS RCT can be highlighted. First, the study provides evidence about the feasibility of routine outcome assessment and feedback.

8.3 Feasibility of the intervention

Since there is a policy commitment to introducing routine outcome assessment (Department of Health, 2001), the financial and human resources needed to implement the intervention warrant consideration.

Substantial effort was put into maximising participation in the intervention, as detailed in Section 6.13. In summary, patient response rate was maximised through involving voluntary sector providers, invitations to FOCUS lunches, prior planning of what to do if a completed postal questionnaire was not received, and a £5 payment for each assessment. Staff response rate was maximised through the FOCUS Local Implementation Group, the provision of FOCUS lunches, publicising the study in health and social services newsletters, presentation of the study at a local research forum, local knowledge of the principal investigator, high visibility of the research team in CMHTs, supplying low-value gifts such as biscuits when recruiting from each CMHT, sending individually signed Christmas cards to participants, postal order gifts for reception and administrative staff, and public acknowledgement of and small payments for staff with several patients in the study.

This level of resources is appropriate for an exploratory trial, but would be difficult to replicate on a wide scale without significant resource implications. Therefore the FOCUS RCT may indicate the maximum response rate that can be obtained within the current culture of care provision, in which outcome measures are not yet a component of routine care.

Unlike traditional outcome studies, the role of the researchers was not to gather all the data via interviews, since the study involved routine assessments and developing a methodology appropriate for mass implementation. Rather, the role of the researchers was to provide the infrastructure support for collection, analysis and feedback which would be needed from a Trust implementing routine outcome assessment. So how much did provision of this infrastructure cost?

Two rounds of the intervention were offered to approximately 100 patients, and baseline and follow-up assessments to 160 patients. Assuming each face-to-face baseline or follow-up assessment takes the same administrative time as a round of the intervention (comprising three postal questionnaires plus one feedback for both staff and patients), this means that (2x100) + (2x160) = 520 rounds of the intervention were offered. Two researchers were each employed for two years, to provide the intervention. This suggests that each researcher can provide 520/4 = 130 rounds of the intervention per year. At 2003 prices, a researcher employed on the RA1B scale at point 6 costs approximately £26,000 per year including associated salary costs. This gives a price per round of intervention offered of approximately £200 (€300). This figure is an under-estimate, since it does not include:

- senior researcher time
- the contributions of other clinical and research staff to the study
- the £15 postal orders sent to patients with the three postal questionnaires
- the opportunity costs of the participants' time – 8.7 minutes for postal questionnaire completion by patients, 7.4 minutes completion time for staff, plus time spent time reading and using the feedback.

8.4 Patient-rated unmet need and quality of life

The relationship between patient-rated unmet need and quality of life was investigated in two ways – using baseline and follow-up assessments for all participants, and using monthly ratings from the intervention group. A mixed picture emerged. Both the absolute baseline level and the change from baseline to follow-up in unmet need were associated with follow-up quality of life. This relationship remained for change scores, but not for absolute values, when controlling for sociodemographic variables and other measures. The cross-sectional time series regression on monthly scores indicated that the absolute value of unmet need one month previously was associated with the current

rating for quality of life. However, the change in unmet need for the previous month was not associated with the change in quality of life for the current month.

Overall there was some evidence of temporal precedence, so hypothesis 3 was partly confirmed. An explanation which is broadly consistent with the data would be that high patient-rated unmet need does reduce quality of life, and when the level of unmet need reduces it takes a period more than one month and less than six months for the full effect to be experienced by the patient in terms of improvements in their quality of life. This would account for changes over a 6-month period being associated with changes in quality of life, whereas absolute values one month previously still contribute to current quality of life. Specifically, Figure 7.3 would suggest that a reduction in patient-rated unmet need in the first month of the study was accompanied by an improvement in quality of life two months later. This result for the first time provides evidence of direction (Bollen, 1989) in the causal relationship, as described in Section 5.4. The implications of this finding for future work are considered in Section 8.6.3.

8.5 Limitations of the FOCUS RCT

Five methodological limitations can be identified: sample selection, outcome measures, data completeness, contamination and blinding. Each will be considered in turn.

8.5.1 Limitation 1: Sample selection

The sample were chosen to be representative of the full range of patients using adult mental health services throughout the London Borough of Croydon. The sample setting was chosen on the basis of its national representativeness for population sociodemographics, as described in Section 6.6. One implication of this approach is that the role of diagnosis was not as prominent as in an efficacy study. The inclusion criteria (listed in Section 6.7) did not include specific diagnostic groups, and the clinical diagnosis was not verified by assessment of a research diagnosis. These choices were made to maximise the external validity for CMHTs, which both include heterogeneous diagnoses on their caseloads and rely on clinical diagnosis. However, since the intervention was not shown to be effective, the sampling approach makes it more difficult to identify whether the intervention was effective for a homogenous diagnostic sub-group.

A future sampling approach could be informed by possible predictors of responsiveness identified in the FOCUS RCT. This approach – theoretical sampling – is a qualitative research technique of initially (as in this study) sampling from a wide range of patients, leading to the generation of hypotheses which are tested with a new and tightly-specified sample in a subsequent study. Patients with a premorbid IQ of more than 100 would be one group to consider for further investigation of the intervention. This is considered further in Section 8.6.1.

8.5.2 Limitation 2: Standardisation of measures

The psychometric properties of MANSA and CANSAS were established using an interviewer to administer the measure. It is therefore possible that their administration through a postal questionnaire may compromise their psychometric properties. Similarly, although CANSAS-P was modified for this study from CANSAS, for which adequate psychometric properties have been demonstrated (Phelan et al, 1995; Andresen, Caputi & Oades, 2000), the psychometric properties for CANSAS-P have not been established. The psychometric consequences for these measures of being self-administered by the patient are amenable to evaluation.

Another possible limitation relates to the sensitivity to change of quality of life. Ratings of quality of life were very stable – the mean rating on a scale from 1 to 7 was 4.25 for both intervention and control groups at baseline, and was 4.27 and 4.20 for these two groups respectively at follow-up. This may indicate an insensitivity to change in quality of life – an undesirable property in a primary outcome measure. There is some evidence for a quality of life 'equilibrium' from the wider literature on quality of life. Cummins (1995) found a uniform population standard of 75% +/- 2.5% of the measurement scale maximum score in 16 general population studies of quality of life. The author notes that this is unlikely to be an artefact of the outcome measure since several different assessments were used. One proposed mechanism is the 'standard drift fallacy', a hypothesis proposing that patients with chronic disorders adjust their standards downwards to reduce the gap between their expectations and their achievements (Katschnig, 1997), leading to higher self-rated quality of life which appears similar to a community sample. However, the mean baseline MANSA score in the FOCUS Study was 54% of the scale maximum: lower than the 75% proposed by Cummins (1995) and not consistent with the standard drift fallacy. More generally, other reviewers have concluded that quality of life changes do occur in general populations, and specifically

shown that changes in quality of life occur for mental health samples in more studies than not (Evans & Huxley, 2002).

8.5.3 Limitation 3: Data completeness

The follow-up data were not complete. Full follow-up data were missing for 18 patients (8 intervention, 10 control), 9 staff (6 intervention, 3 control) and 2 casenote audits (1 intervention, 1 control). There was no apparent difference between patients with and without full follow-up data. To minimise the impact of selective attrition, all analysis was undertaken on an intention-to-treat basis, followed by a sensitivity check substituting baseline for follow-up scores for these 18 patients. However, the possibility that the missing patients had a systematically different response to the intervention, which would have altered the findings had their follow-up data been included, cannot be discounted.

8.5.4 Limitation 4: Contamination

The possibility of contamination between the intervention and the control group was noted in Section 8.2.2. Contamination in this context indicates that the anticipated behavioural changes by staff – more reflective practice leading to care plans focused on desirable outcomes – may also have been present in the control group. Clustering is the best solution to investigate behavioural interventions designed to change clinical practice (Gilbody & Whitty, 2002), and would be appropriate for investigating routine outcome assessment and feedback (Gilbody et al, 2001). In this context, the appropriate cluster would be individual staff, since any further dependence (*e.g.* at the CMHT or site level) is likely to be of less significance. This approach will introduce an intra-cluster correlation, since patients seen by each member of staff will be more similar to each other than to patients seen by other staff. Conventional sample size calculations assume that each observation is independent, and failing to account for an intra-cluster correlation (a 'unit-of-analysis error') leads to an over-estimate of significance and increased likelihood of a Type 1 (false positive) error. Therefore an implication of this design change is that the sample size would need to be increased.

8.5.5 Limitation 5: Blinding

The final limitation was that blinding was only minimally maintained. Neither patients nor staff were blind to allocation status. The analysis was also not undertaken on a blind

basis. To minimise bias in the analysis phase, the analytical strategy was agreed with an independent statistician before looking at the follow-up data.

The researchers conducting the follow-up interviews were partially blind – they were unable to guess allocation status for 81 (57%) of 143 staff interviews and for 41 (29%) of 140 patient interviews. However, where they did guess allocation status they were correct for 97 (92%) of their 105 intervention group ratings, and for 53 (95%) of their 56 control group ratings. How could the estimate of blindness be improved? A better approach would have been to require the researcher to guess as to allocation status, rather than allowing a 'Don't know' response.

Blindness is a recognised difficulty in evaluating complex health interventions (Medical Research Council, 2000), and in the FOCUS Study would be difficult to improve substantially. Blinding could have been marginally increased in two ways. First, by ensuring the follow-up interviews were undertaken by a different researcher group to those involved in recruitment into and running of the trial. In the FOCUS Study, however, many patients (despite requests) did unblind themselves during the follow-up interview. Therefore the potential gains of using a different researcher group would need to be weighed against the resource implications, to ensure the costs are not disproportionate for a marginally increased level of blindness. Second, analysis by an independent statistician would improve blindness, although they would need to be unaware of whether the larger (n=101) or smaller (n=59) group was the intervention group.

Blinding is especially important when patient preferences are likely to differ between treatments. In simple health interventions, such as a comparison of two types of antidepressant which are indistinguishable to the patient, there may not be a strong patient preference. In complex health interventions, by contrast, patients are likely to have views about the relative merits of each treatment arm, since the implications for them of their allocation status are more apparent. The lack of consideration for patient preference in standard RCT methodology has led some commentators to conclude that until the issue is addressed, "*we might never know, even approximately, how much of modern medicine is attributable to psychological processes*" (McPherson, Britton & Wennberg, 1997, p. 652). In response to this concern, a variant on the RCT design has been developed – the **patient preference trial** (Torgerson & Sibbald, 1998). In this

variant, patients with treatment preferences are allocated to their desired treatment, and only those who do not have strong views are randomised conventionally. However, the analytical strategy for such trials is unclear, and they are likely to be larger and cost more than traditional RCTs.

A further methodological advance is also needed, which is not normally considered (*e.g.* Chalmers, 2001). In complex health interventions, it is more likely that staff will provide, receive or evaluate the intervention. Since single blinding (*i.e.* of participants) is often impossible, the preferences of staff may be as important as those of patients. For example, in the FOCUS Study both the staff and the patient received the intervention. Although the intention in any clinical trial is that genuine clinical equipoise (*i.e.* no staff preference) exists regarding treatment effectiveness, this is often difficult to achieve in practice. It was notable that when discussing the intervention with staff prior to the start of the FOCUS Study, both strongly positive and strongly negative views were expressed, especially in relation to the feedback of the HAS data. Patient preference designs neglect the possibility of staff preferences, and further methodological developments are needed. One approach is described in Section 9.2.3.

8.6 Future work

Three strands of future research are indicated by the results of this study: developing the finding that higher premorbid IQ was associated with more benefit from the intervention, amending and re-testing the FOCUS Model, and testing whether providing care on the basis of need leads to improvements in quality of life.

8.6.1 Future work 1: Premorbid IQ

The first potential avenue of future work would develop the premorbid IQ findings. The FOCUS RCT indicated the potential to benefit from the intervention for those with a premorbid IQ of 111 or more, and also provided indications of a weaker effect (not detected due to lack of power) for those with premorbid IQ of 100 or more. The number of patients in this sub-group reduces power and so these results can only be hypothesis-generating. However, it is plausible to suggest that a larger sample size would have found a significant effect of the intervention on both patient-rated unmet need and quality of life for patients with higher premorbid IQs.

Why should patients with higher premorbid IQ differentially benefit from the intervention? A first hypothesis is that more intellectually able patients may make better use of the intervention, both in re-framing their own experience and in being prompted to discuss with staff their care plan or the therapeutic alliance. More importantly, staff may find it easier to use the feedback to inform future care with more intellectually able patients. This would be consistent with the results, which found more alteration in both staff and patient behaviour in the higher premorbid IQ sub-group (shown in Table 7.12) than in the whole sample (shown in Table 7.10). This hypothesis could be investigated through replicating the FOCUS RCT, with an added inclusion criterion of the patient's premorbid IQ being 100 or more. Given the preliminary evidence of effectiveness for this patient population from the FOCUS RCT, a more detailed evaluation assessing both effectiveness and cost-effectiveness would be warranted.

A second hypothesis is that lower premorbid IQ patients were disadvantaged by the written medium of the feedback, and require other means of information transfer. The pilot study described in Section 6.12 investigated the best means of presenting the feedback, but the patient numbers involved were small (n=5) and differing preferences were expressed. The combined graphical and text-based feedback which was developed on the basis of this feedback may not have been the most accessible for patients with lower premorbid IQ. This hypothesis is not supported by the data presented in Table 7.10 in which 84% of patients stated they understood the feedback, although as noted in Section 8.2.5 this self-report data may have been inflated by social desirability bias – the patient being too embarrassed to admit they did not understand the feedback.

This hypothesis could be tested by further pilot study work to compare different modes of information transfer for how comprehensible they are to patients of different intellectual abilities. For example, written feedback could be compared with a face-to-face meeting between the patient, the staff and the feedback provider for high and low premorbid IQ patients.

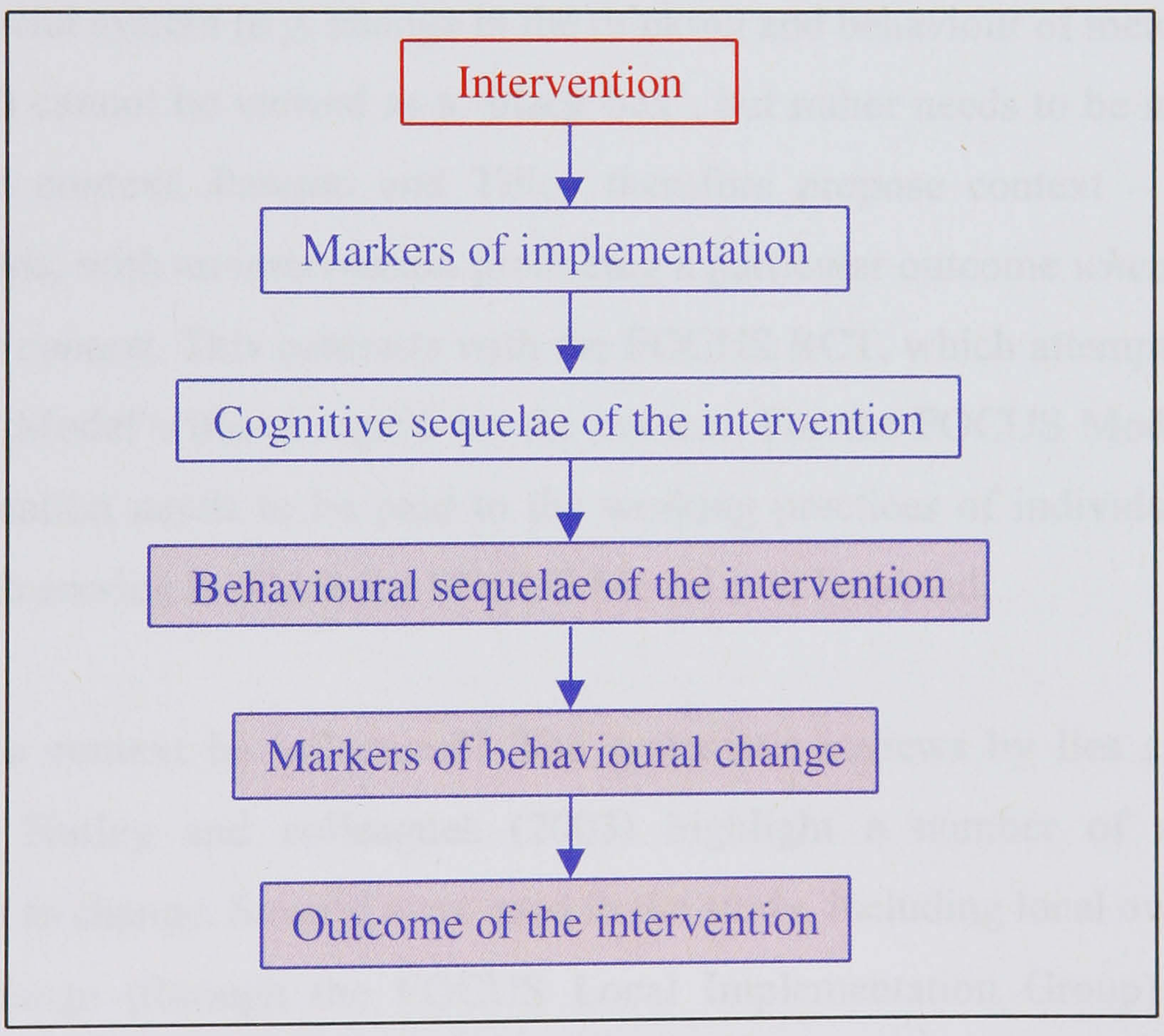
A third hypothesis is that lower IQ was associated with a more coercive existing style of interaction, which was less amenable to shaping through the intervention. Investigating this hypothesis would involve:

1. measuring baseline levels of coercion and associated variables (such as empowerment)
2. monitoring the interactions (especially regarding the assessments and feedback) which occur between patient and staff during the study using conversation analysis techniques (Ten Have, 1999)
3. analysing whether baseline levels or style of interaction predict responsiveness to the intervention, independent of cognitive measures.

8.6.2 Future work 2: Amending the FOCUS Model

As described in Section 8.2.5, evidence from the FOCUS RCT validates some components of the FOCUS Model and does not validate others. The FOCUS Model can be separated into six elements, shown in Figure 8.3, with non-validated elements in greyed boxes.

Figure 8.3: Summary of FOCUS Model



This analysis indicates that the FOCUS Model needs to be modified to more strongly encourage changes in the behaviour of staff and patients. Since most literature relates to changing staff behaviour, and the preliminary data from the FOCUS RCT indicated that (self-reported) staff behaviour change was more likely than patient behaviour change, the focus will be on the staff version of the FOCUS Model.

Several approaches to changing staff behaviour can be explored, based on the Green and Eriksen (1988) model described in Section 6.13. This will involve increasing the predisposing awareness of the need for behaviour change, and enabling the behaviour change more than was achieved in the FOCUS RCT. Two recent systematic reviews have recommended a range of approaches to encouraging practice change (Iles & Sutherland, 2001; Nutley, Percy-Smith & Solesbury, 2003). The recommendations from these two reviews will be considered in terms of predisposing (Section 8.6.2.1) and enabling (Section 8.6.2.2) change in staff practice.

8.6.2.1 Increasing the predisposition to behaviour change

More attention needs to be given to the context of the intervention, since a service whose shared beliefs are congruent with the use of outcome measures is necessary if the intervention is not to be 'swimming against the tide'. The role of context is considered more in the *Realistic Evaluation* approach to evaluating social programmes advocated by Pawson and Tilley (1997). This approach highlights that any intervention whose target is a social system (*e.g.* change in the thinking and behaviour of mental health staff and patients) cannot be viewed as a 'black box', but rather needs to be understood and evaluated in context. Pawson and Tilley therefore propose context – mechanism – outcome triads, with an intervention producing a particular outcome *when it is provided in a specific context*. This contrasts with the FOCUS RCT, which attempted to evaluate the FOCUS Model without regard for the context. For the FOCUS Model, this means that consideration needs to be paid to the working practices of individual staff in the mental health service in which the FOCUS Model is to be tested.

How can the context be influenced? The systematic reviews by Iles and Sutherland (2001) and Nutley and colleagues (2003) highlight a number of approaches to predisposing to change. Several were used in the study, including local ownership of the proposed change (through the FOCUS Local Implementation Group) and personal contacts between and co-location of researchers and practitioners (FOCUS researchers were highly visible in Croydon, and worked in the teams). However, the reviews also highlight the importance of organisational beliefs and working practices, the need for research programmes rather than isolated research studies, and (perhaps most relevantly) for demonstration sites. A demonstration site would be a service which uses outcome measures as a routine element of care on an ongoing basis. This contrasts with the participating CMHTs in the FOCUS RCT, which did provide and to some extent use

the information during the study, but then ceased to use the outcome information after the trial was completed.

How could such a demonstration site be developed, with outcome-focussed organisational beliefs and working practices? Specifically, how can the use of outcome data be moved from an activity peripheral to clinical work (*i.e.* participation in the FOCUS RCT) to being a central component? Several approaches have been suggested (Andrews et al, 1994; Slade et al, 1999d):

- Providing care on the basis of a standardised assessment of need – this is discussed further in Section 8.6.3
- The routine use of clinical guidelines to identify and inform the choice of treatment options, accompanied by methods (*e.g.* clinical supervision, audit) to ensure guidelines are followed, as described in Section 1.3
- The integration of outcome information into the existing Care Programme Approach care planning process. This would involve the imposition of targets for reviewing care plans, including explicit consideration of whether to continue, alter or discontinue treatment.
- The development of both paper-and-pencil and computer administered versions of the chosen outcome measures, and the identification of optimal implementation approaches (*e.g.* by the patient before their appointment using a computer in the waiting room). The goal would be to establish a context in which "*routine administration [of outcome measures] is seen as a natural part of clinical practice, both by consumer and by clinician*" (Andrews et al, 1994, p.69).
- The introduction of a system of rewards and sanctions based on local performance in using outcome measures – there need to be tangible gains for staff if they are to change their practice
- The development and use of an outcomes database, with information supplied (with minimal effort for staff) both to inform clinical care and to influence service development, as per Section 1.3. Specifically, resource allocation decisions should be explicitly linked to and based on outcome information.
- The minimisation of other practice changes.

Overall, the development of a service which actively uses outcome information in care planning and reviewing would provide a more appropriate test-bed in which to evaluate

routine outcome assessment. It would add credibility to the research and have high salience for other staff. This would predispose staff to seeing outcome information as central to, rather than peripheral to, the provision of high-quality mental health care.

8.6.2.2 Enabling behaviour change

In addition to developing a demonstration site in which the intervention of assessment plus feedback is more congruent with organisational beliefs, it may be possible to improve the intervention. The systematic reviews of practice change approaches also recommended approaches relevant to enabling practice change (Iles & Sutherland, 2001; Nutley et al, 2003). Again, some approaches were used in the study – frequent reinforcement through reminders (the postal questionnaires and feedback), tailored material for target audiences (each feedback was specific to the patient), and use of incentives (FOCUS lunches, postal orders).

The intervention comprises two elements: assessment plus feedback. There was no evidence that the choice of outcome measures for routine assessment could be improved. However, the feedback could be improved in two ways.

First, the postal questionnaires could be more focussed on identifying topics which the patient would like to discuss with staff. This approach is used in the 2-COM questionnaire (van Os, Altamura, Bobes et al, 2002). The 2-COM questionnaire is a 19-domain patient-rated measure based on the CANSAS domains, which asks for each domain "Is this a problem for you?" and "Would you like to talk about it?". This not only allows the identification of unmet needs, but also the patient's prioritisation of which domains are most important for them. A randomised controlled trial investigated using versus not using 2-COM with 134 patients with schizophrenia (van Os, Altamura, Bobes et al, 2004). Beneficial effects on care plan (OR=3.7) and patient-reported quality of patient-clinician communication (B=0.33, p=0.03) were found. However, using 2-COM was associated with an increase of 13 minutes in consultation time – basing care more on the individual patient's preference may decrease the capacity of the service. This provides some indications that identifying topics of most relevance to the patient in the feedback may beneficially impact on the process of care.

Secondly, the feedback could more actively promote behavioural change. The feedback in this study gave a summary of results with areas of disagreement highlighted. More

prescriptive advice for action could also be given, both in relation to process and content of care. This might involve suggestions for specific interventions or referrals, when a particular need is identified.

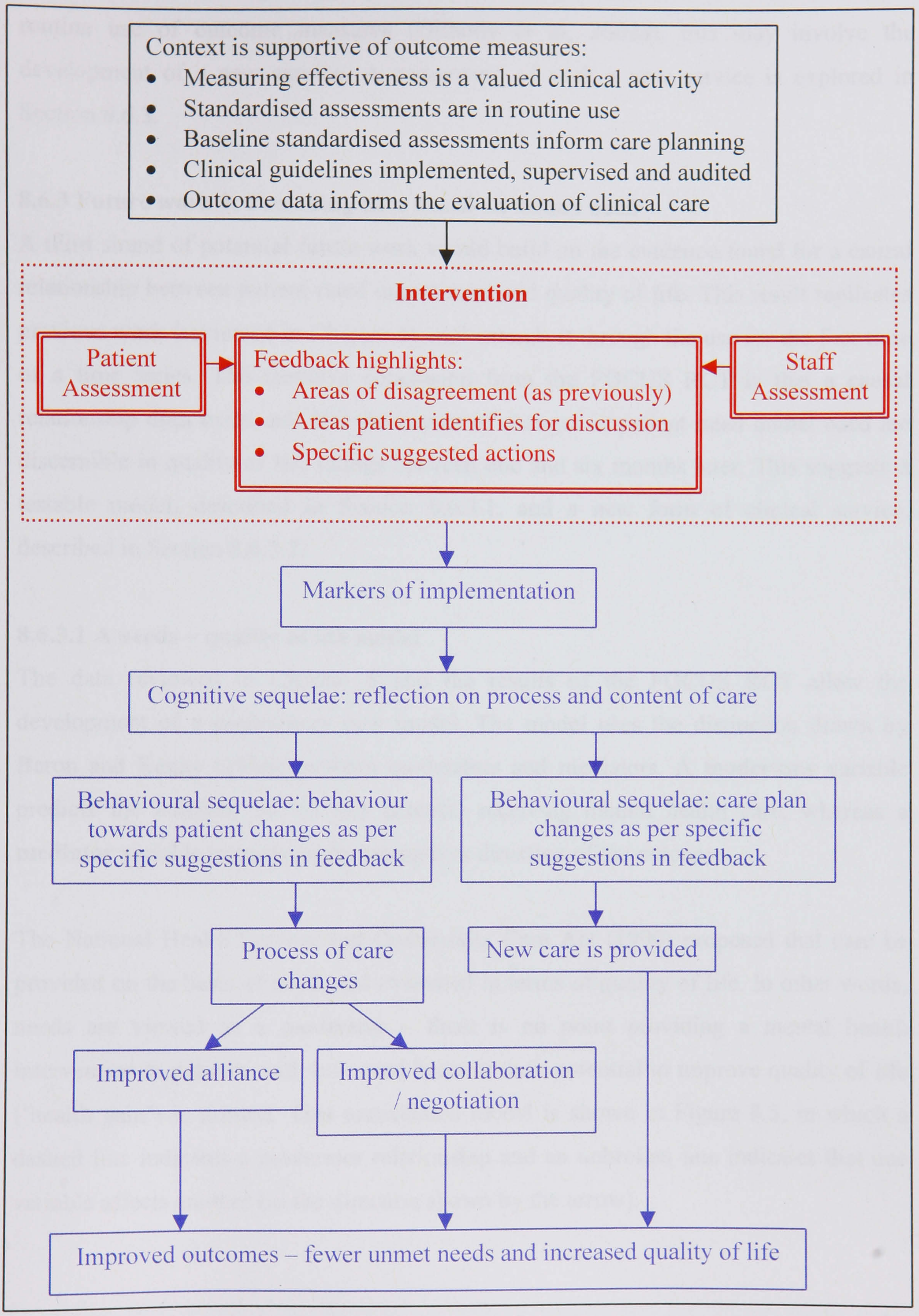
Evans and colleagues distinguish between a 'need' – which can arise for many reasons, and have many possible solutions, not all of which will be interventions – and a 'need for care' – the subset of needs which can be met through interventions (Evans, Greenhalgh & Connelly, 2000). The development of intervention options for each need identified by CANSAS would require validation, since not all needs can be converted into a need for care. This approach of listing intervention options is used by some other approach to needs assessment, such as the Medical Research Council Need for Care Assessment Schedule (Brewin et al, 1987) and the Cardinal Needs Schedule (CNS) (Marshall, Hogg, Gath et al, 1995). The impact on symptomatology, social functioning and level of unmet need of an intervention called 'needs feedback' – feeding back the CNS assessment to community psychiatric nurses – is currently being evaluated (Lockwood & Marshall, 1999).

However, a randomised controlled trial using the Camberwell Assessment of Need for the Elderly (CANE) (Reynolds, Thornicroft, Abas et al, 2000) investigated feedback of needs assessment to staff at a psychiatric day hospital for older people (Ashaye, Livingston & Orrell, 2003). This study found no difference in unmet need, social disability or behavioural functioning between the 58 control group patients (who received standard CPA) and the 54 intervention group patients whose staff received feedback on the CANE results. The authors concluded that using a standardised assessment of need offered no advantages over the careful unstandardised assessment of need which was part of routine care. This contrasts with the findings of earlier, uncontrolled studies (Macpherson, Jerrom, Lott et al, 1999). Nonetheless, it may be worth including specific care recommendations in the feedback, which are based on the outcome assessments, with the aim of beneficially impacting on the content of care.

8.6.2.3 The amended FOCUS Model

The discussion in Sections 8.6.2.1 and 8.6.2.2 indicates that more attention needs to be given to the context of the intervention, and the feedback needs to be more tailored towards promoting specific behavioural responses. The staff version of the amended FOCUS Model is shown in Figure 8.4.

Figure 8.4: The amended FOCUS Model



The patient version of the amended FOCUS Model differs only in the changing of "*towards patient*" to "*towards staff*" for the two behavioural sequelae.

The amended FOCUS Model can be tested using a similar design and hypotheses to the FOCUS RCT, providing an appropriate context is established. Given the current lack of routine use of outcome measures (Gilbody et al, 2002a), this may involve the development of a new service. A component of such a new service is explored in Section 8.6.3.

8.6.3 Future work 3: Providing care based on unmet need

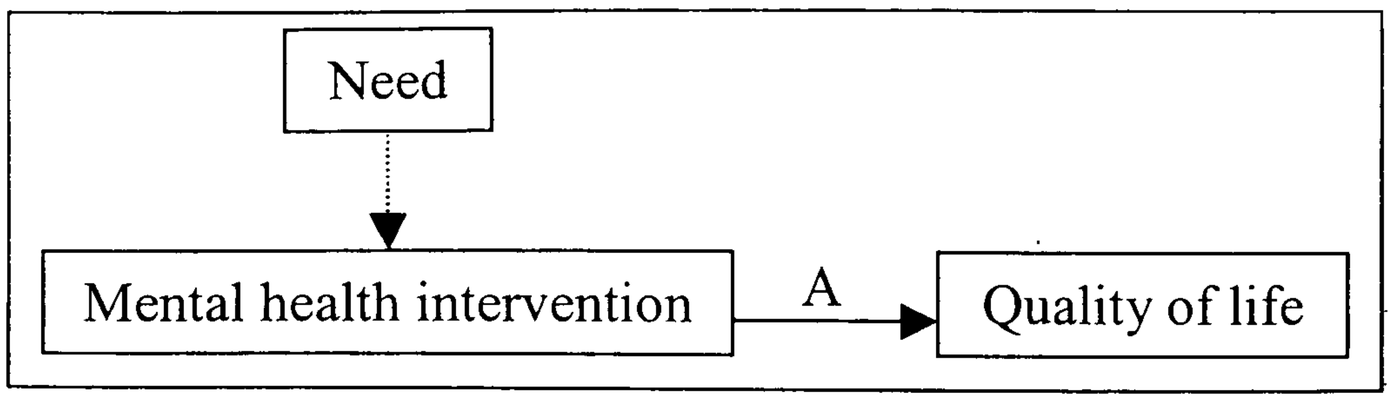
A third strand of potential future work would build on the evidence found for a causal relationship between patient-rated unmet need and quality of life. This result replicates previous work (reviewed in Chapter 4), and extends it through the use for the first time of a time series. The tentative conclusion from the FOCUS RCT is that a causal relationship does exist, and that the impact of changes in patient-rated unmet need are discernible in quality of life ratings between one and six months later. This suggests a testable model, described in Section 8.6.3.1, and a new form of clinical service, described in Section 8.6.3.2.

8.6.3.1 A needs – quality of life model

The data reviewed in Chapter 5 and the results of the FOCUS RCT allow the development of a preliminary new model. The model uses the distinction drawn by Baron and Kenny (1986) between moderators and mediators. A **moderator** variable predicts the outcome of (in this context) receiving mental health care, whereas a **mediator** variable impacts on the strength or direction of the outcome.

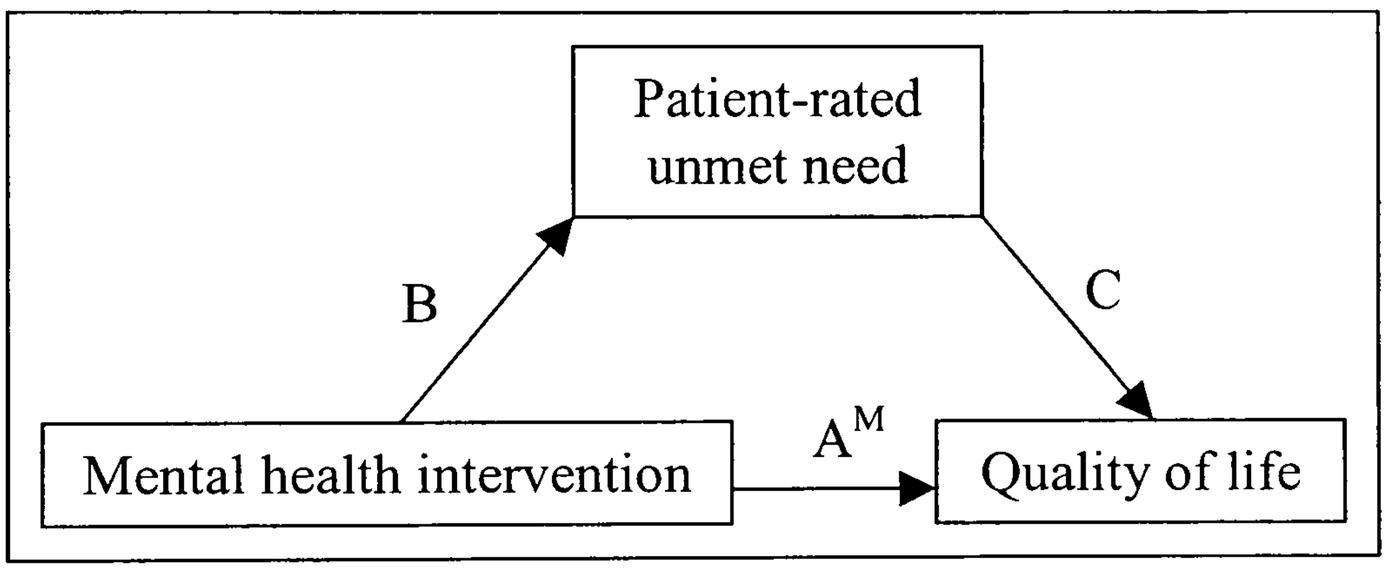
The National Health Service and Community Care Act (1990) proposed that care be provided on the basis of need, and evaluated in terms of quality of life. In other words, needs are viewed as a moderator – there is no point providing a mental health intervention to patients with low need because their potential to improve quality of life ('health gain') is limited. This unmediated model is shown in Figure 8.5, in which a dashed line indicates a moderator relationship and an unbroken line indicates that one variable affects another (in the direction shown by the arrow).

Figure 8.5: Unmediated model of the relationship between a mental health intervention and quality of life



The review in Chapter 5 indicates that the most appropriate aspect of 'need' to consider is patient-rated unmet need. The results of the analysis in Chapter 5 and the FOCUS RCT indicate that patient-rated unmet need may, as well as being a moderator, also be a mediator. The model in which patient-rated unmet need is a mediator is shown in Figure 8.6.

Figure 8.6: Mediated model of the relationship between a mental health intervention and quality of life



In this model, the effect of the mental health intervention is mediated by its impact on patient-rated unmet need. If the mental health intervention no longer affects quality of life after patient-rated unmet need has been controlled for (and so path A^M is zero), then this is called complete mediation. If the path from mental health intervention to quality of life is reduced in absolute size but still different from zero when patient-rated unmet need is controlled for, this is called partial mediation.

Testing the mediated model shown in Figure 8.6 involves four steps (Baron and Kenny, 1986):

1. Show that the mental health intervention is correlated with quality of life, by using quality of life as the dependent variable in a regression equation and the intervention (*i.e.* allocation status) as the independent variable. In other words, estimate and test path A in Figure 8.5, to establish that there is an effect that may be mediated.

2. Show that the mental health intervention is correlated with patient-rated unmet need, by using patient-rated unmet need as the dependent variable in a regression equation and the intervention as the independent variable. This will estimate and test path B in Figure 8.6.
3. Show that patient-rated unmet need affects quality of life, by using quality of life as the dependent variable in a regression equation and both the intervention and patient-rated unmet need as independent variables. This will estimate and test path C in Figure 8.6. It is not sufficient just to correlate patient-rated unmet need with quality of life, since they may be correlated because they are both caused by the intervention – the intervention must be controlled for in establishing the effect of patient-rated unmet need on quality of life.
4. To establish that patient-rated unmet need completely mediates the intervention – quality of life relationship, the effect of the intervention on quality of life controlling for patient-rated unmet need should be zero. This will estimate and test path A^M in Figure 8.6. The effects in steps 3 and 4 are estimated in the same regression equation.

Step 1 was tested in Section 7.6, which did not show that the intervention being tested in the FOCUS RCT was associated with quality of life. Therefore further development of the model cannot be investigated empirically using data from the FOCUS RCT, since the intervention did not influence quality of life. The remaining discussion in this section is therefore hypothesis-generating.

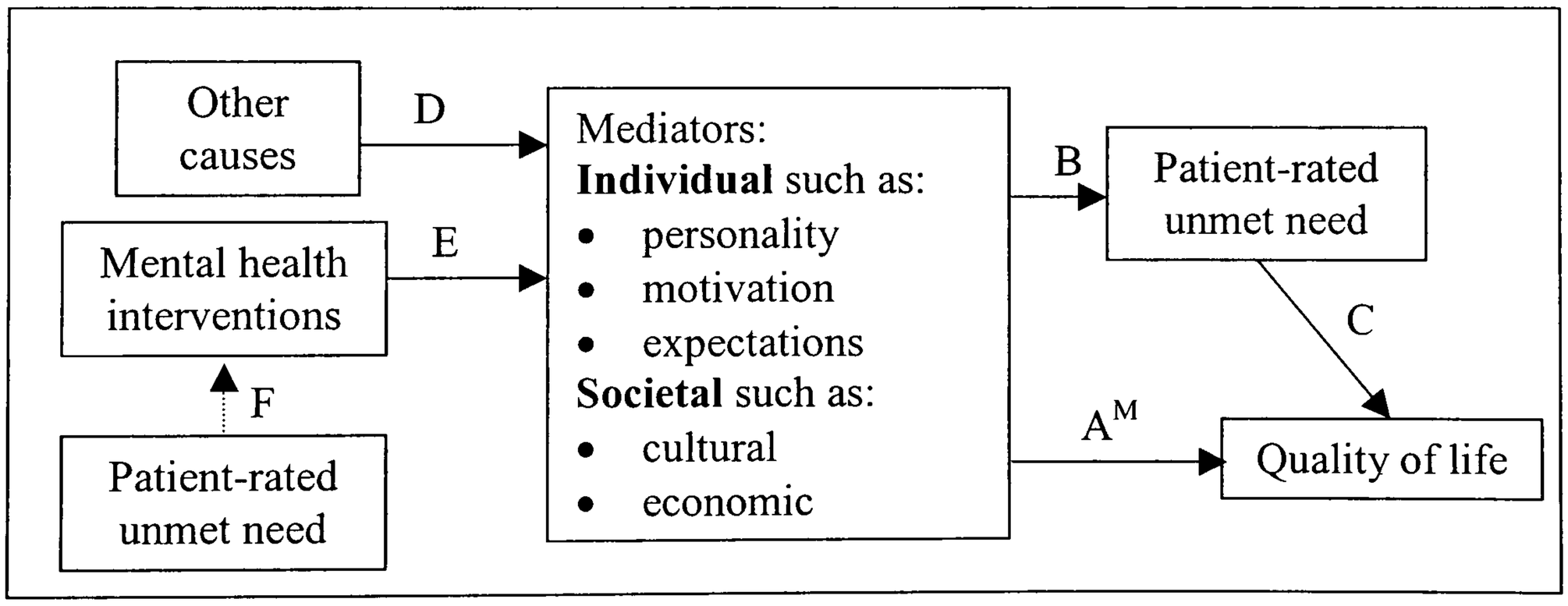
What is the strength of path A in Figure 8.5? In other words, how strongly do interventions impact on quality of life? In the retrospective re-analysis of existing data from South Verona reported in Section 5.6, the magnitude of association found in this study was relatively modest – meeting one unmet need would lead to a one percent change in subjective quality of life (with the 95% confidence interval indicating the maximum change consistent with the data as 2%).

This finding that mental health care has a limited impact on quality of life is consistent with results from other areas of medicine, such as cardiac care (Beck, Joseph, Belisle et al, 2002). This indicates that consideration of a broader range of influences than solely health care interventions will be needed, for a more complete understanding of how to influence quality of life. Two levels of influence can be distinguished: causes and

mediators. Potential causes include all other life events and experiences apart from mental health interventions which happen to a patient and impact on their quality of life. This is illustrated by the impact of falling in love on both patient-rated unmet need and quality of life. Mediators can be individual or societal. Examples of individual mediators shown to influence quality of life include personality (Ruggeri et al, 2001), motivation (Grahn, Ekdahl & Borgquist, 2002) and expectations (Ruggeri & Dall'Agnola, 1993). Societal mediators shown to influence quality of life include culture (Michalos & Zumbo, 2001) and economic prosperity (Schyns, 1998).

Patient-rated unmet need has emerged as both a moderator and a mediator for the relationship between an intervention and quality of life. Furthermore, the need to consider other causes of quality of life is empirically indicated by the relatively weak influence of health care interventions on quality of life. A summary model is shown in Figure 8.7, in which the "Other causes" box encompasses all life events and experiences other than mental health interventions which impact on quality of life.

Figure 8.7: General model of the relationship between mental health interventions and quality of life



This preliminary model is intended to be hypothesis-generating, and future investigation may lead to several changes. For example, if path D in Figure 8.7 is much stronger than path E, then what is viewed as an 'intervention' may need to become broader to encompass other causal influences on quality of life. In practice, this would mean that if quality of life is improved for the general population by having more money, playing sports and going on holiday, then the routine inclusion within a multidisciplinary mental health team of staff with expertise in welfare benefits advice, sports coaching and holiday planning should be considered.

As another example, mediators may have a differential impact on different causes. For instance, the impact of expectations on the effectiveness of psychological therapy may not be the same as the impact of expectations on the effectiveness of pharmacotherapy. This would indicate that the 'Mediators' box needs to be disaggregated.

As a third example, the individual and societal mediators are assumed in this model to influence the patient's appraisal of their unmet needs. Specifically, the distinction between an unmet need (*i.e.* a current serious problem) and a met need (*i.e.* no problem or only a moderate problem) will be influenced by psychological characteristics (*e.g.* some patients minimise and some maximise their problems) and contextual characteristics (*e.g.* based on the reference group with which the patient compares him or herself). However, it may be the impact of unmet needs on quality of life that is more strongly influenced by these mediators. In this case, the 'Mediators' box should also feature in path C.

8.6.3.2 A new type of adult mental health service

If the causal relationship between patient-rated unmet need and quality of life is confirmed through replication, then this suggests an alternative service model. The characteristics of such a service would be a focus on the patient's perspective in assessment, the systematic identification of the full range of health and social care needs of the patient, the development of innovative services to address these needs, and the evaluation of the success of the service in terms of impact on quality of life. This contrasts with the dominant current service model, which can be characterised as having an emphasis on the professional perspective, the non-systematic assessment of needs by a range of professionals, a focus on symptom relief and risk reduction, and disagreement about how to evaluate the outcome of care. This links with the discussion in Section 3.8.2 regarding patient-defined versus professionally defined approaches to understanding mental problems. The new service would need to be based on a testable model of the intended effects of increased user involvement, and such a model is currently lacking (Crawford, Rutter, Manley et al, 2002).

This chapter has discussed the results of the FOCUS RCT. The concluding chapter will summarise the scientific, clinical and policy implications of the FOCUS Study as a whole.

Chapter 9

Conclusion

9.1 Evaluation of a complex health intervention

An inter-connected series of research studies have been reported, covering the first three stages of the MRC Framework for Complex Interventions to Improve Health (Campbell et al, 2000). A systematic review of outcome domains and routine outcome assessment principles in Chapter 2 and 3 informed the development of a testable model in Chapter 4. One aspect of the FOCUS Model – the relationship between the two primary outcomes – was investigated through a hypothesis-driven retrospective re-analysis of existing data in Chapter 5. This theory – model – theory iteration is consistent with the MRC Framework: '*Progression from one phase to another may not be linear. In many cases an iterative process occurs*' (Campbell et al, 2000, p. 694). The FOCUS Model was then evaluated using an exploratory randomised controlled trial, described in Chapters 6 and 7. An amended version of the FOCUS Model was proposed in Chapter 8.

The FOCUS Model is complex, with several potentially active ingredients. This is illustrated by the reports from patients and staff on their experience of participating in the FOCUS RCT, assessed using the Impact of Involvement form. Patients were asked about their experiences of filling in the postal questionnaire. All quotes are verbatim:

"Brought up things I wouldn't normally talk about."

"Made me realise that I had a voice."

"Highlighted how our relationship has improved."

"Made me feel closer to the service, more attention. Like a session with the doctor through the post."

"Made me think of what kind of care I was getting and whether it was any good for me."

"I wanted to know whether he shared my view about the treatment I was getting."

"Made me think about appointeeship, regarding control over money – this has caused a few rows – made me more aware of what staff should be doing."

"Made me think he didn't bother with me."

Patients were also asked about receiving the feedback:

"I feel it made me stop and evaluate my situation and self."

"It gave me a perspective on both sides."

"Made me realize we are on the same wavelength."

"Made me realise how angry I was with him, and how I worked it out. Was good for us."

"I looked at the different answers and it made me feel more believed because my keyworker say the same problems that I did."

"Made me think I wasn't getting as much care as I should be getting."

"I wondered how he was filling in questionnaire without seeing me."

"I felt she assumed I was progressing and coping better than I actually was."

Staff were asked about the impact of filling in the postal questionnaires:

"See from Service User perspective."

"At the beginning made sure addressing the area FOCUS identified."

"Consider his unmet needs, and try to resolve as much as I could."

"It helped me to assess whether my delivery of care is holistic."

"Reflection on clinical practice and difficulty I have encountered at times with his demands."

"Differences between my view and her view of her care."

"Made me reflect on the important relationships in her life."

"Made me question a lot of things. Have to answer questions honestly. Made me think."

"I will now question my clients more vigorously about what I do that is effective and what can be changed or developed."

"Issues important to patients, e.g. sexual relationships, accommodation etc. are not always explored."

"Only [had an impact] in our therapeutic relationship, but not in any other way – a lot of it was irrelevant."

Staff were also asked about receiving the feedback:

> *"Sometimes surprises and differences in interpretations, mine and service user's. Checked it out with her."*
>
> *"We talked about it and the discrepancies shown, which was useful."*
>
> *"It made me think of what is needed."*
>
> *"Whether I should be looking at things I can do to make our relationship better."*
>
> *"Looking for ways to improve their care based upon what they say."*
>
> *"Feedback provided me with details of how the client viewed our relationship as compared to myself. This subsequently made me more focused during our meetings and the discussions we had."*

The FOCUS Study developed and evaluated a complex intervention, and as noted in Section 1.1 was intended to have scientific, clinical and policy implications. This concluding chapter summarises the implications of the FOCUS Study in each of these three areas.

9.2 Scientific implications

Four main scientific conclusions can be identified, relating to the scientific framework, the FOCUS Model, the methodology and the choice of outcome.

9.2.1 Implication 1: Scientific framework

The FOCUS Study has followed a scientific framework – the MRC Framework for Evaluating Complex Health Interventions. The placing of each element within the MRC Framework is shown in Figure 9.1, with arrows indicating the direction of development of scientific knowledge.

Figure 9.1: Relationship between the FOCUS Study and the MRC Framework

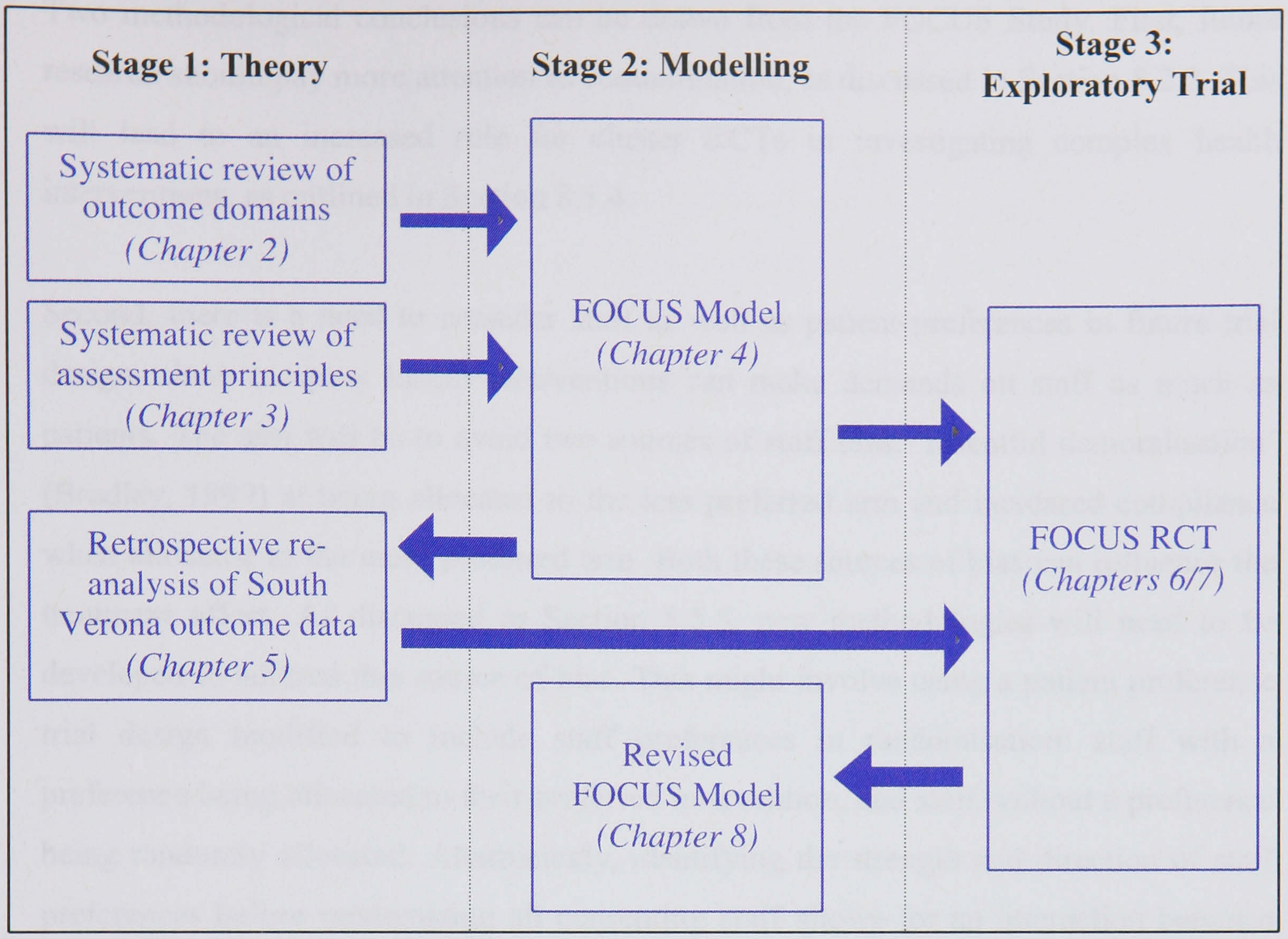


Figure 9.1 illustrates that the MRC Framework was applicable to this complex health intervention. It fostered the development of a theory-based and falsifiable model, which was then tested. Consideration of the results both allowed a simplified version of the original FOCUS Model to be developed (Figure 8.3), and indicated aspects of the model which were validated (Section 8.2.5). This informed the development of a new model in Section 8.6.2.3. The MRC Framework is to be commended as an approach to investigating complex health interventions.

9.2.2 Implication 2: The FOCUS Model

The FOCUS RCT indicates the need to give more attention in the FOCUS Model to promoting behaviour change in future research. Specifically, the 'bolting on' of such a complex intervention to an existing service may be less effective than evaluating the impact of the intervention within a service which is explicitly developed to have a focus on outcomes. Section 8.6.2 considered this issue in more detail, and concluded that changes were indicated both in the content of the feedback and, more fundamentally, the context in which the intervention was provided. An amended version of the FOCUS Model was proposed in Section 8.6.2.3.

9.2.3 Implication 3: Methodology

Two methodological conclusions can be drawn from the FOCUS Study. First, future research should pay more attention to contamination, as discussed in Section 8.2.2. This will lead to an increased role for cluster RCTs in investigating complex health interventions, as outlined in Section 8.5.4.

Second, there is a need to consider staff as well as patient preferences in future trial design, since complex health interventions can make demands on staff as much as patients. The aim will be to avoid two sources of staff bias: 'resentful demoralisation' (Bradley, 1999) at being allocated to the less preferred arm and increased compliance when allocated to the more preferred arm. Both these sources of bias can influence the treatment effect. As discussed in Section 8.5.5, new methodologies will need to be developed to address this source of bias. This might involve using a patient preference trial design modified to include staff preferences in randomisation: staff with a preference being allocated to their preferred intervention, and staff without a preference being randomly allocated. Alternatively, identifying the strength and direction of staff preferences before randomising all consenting staff allows for an interaction between staff preference and patient outcome to be assessed.

9.2.4 Implication 4: Outcome measure

Quality of life is an important outcome for people using mental health service. However, as discussed in Section 8.5.2, it may be insufficiently sensitive to change for use as a primary outcome. This may not be due to insensitive measurement – Section 8.6.3.1 argued that it may be because the effectiveness of mental health interventions is insufficient to substantially impact on quality of life.

Quality of life is a distal outcome, impacted on by many factors other than mental health care. Evaluation of mental health services may need to use more proximal measures, whilst avoiding outcomes such as admission rates which have low salience for people using mental health services. A candidate outcome is patient-rated unmet need, discussed further in Section 9.3.3. Other proximal outcome domains to consider in any replication of the FOCUS RCT include empowerment, perceived involvement, and satisfaction with care.

9.3 Clinical implications

Three clinical implications arise from the FOCUS RCT, relating to changing behaviour, premorbid IQ and patient-rated unmet need.

9.3.1 Implication 5: Changing staff and patient behaviour

Staff and patient behaviour did not change from receiving the intervention. Experience from implementing other service-level changes may be relevant. The Care Programme Approach, intended to ensure effective care planning for patients with more severe mental health problems, was introduced in 1993. A national survey published in 1999 indicated that its implementation was inconsistent (Bindman, Beck, Glover et al, 1999), and uneven implementation was again found in 2003 (Simpson, Miller & Bowers, 2003). This difficulty in implementing a change on a national level indicates that care practices alter only slowly over time.

More generally, although more prescriptive clinical guidelines are being disseminated by the National Institute for Clinical Excellence, in isolation these will probably only incrementally impact on practice. A more multi-level and tailored approach is required (Grol & Grimshaw, 2003). Specific recommendations on ways of maximising the impact of routine outcome assessment were made in Sections 8.6.2.1 and 8.6.2.2, including more prescriptive and personalised feedback, the increased use of clinical guidelines (supported by routine clinical supervision and audited regularly), and the development of a demonstration site in which routine assessment and use of outcome information using standardised measures is a valued clinical activity.

9.3.2 Implication 6: Premorbid IQ

The intervention was definitely effective for the top quarter of premorbid IQ patients, and possibly for the top half. The role of cognitive variables as predictors of intervention effectiveness requires further elaboration, as noted in Section 8.2.4. In Section 8.6.1 it was suggested that the FOCUS RCT does provide support for a future randomised controlled trial, using theoretical sampling (described in Section 8.5.1) to investigate the intervention with higher premorbid IQ patients. Based on the results from the FOCUS RCT, such a study would require a sample of 74 patients in each group (148 in total) to detect a difference of 54% of improvers in quality of life in the intervention group compared with 30% in the control group (*i.e.* the ratio reported in Section 7.7) with power of 80% and $p<0.05$.

9.3.3 Implication 7: Patient-rated unmet need

The third clinical implication arises from the evidence for a causal relationship between patient-rated unmet need and quality of life. If there is a casual relationship between these two outcomes, then baseline assessment of need has two purposes. First, level of patient-rated unmet need may help identify the sub-group for whom substantial change in quality of life is an appropriate goal. This group should be prioritised if maximising quality of life is the goal of the service. Second, an identified unmet need is a target for change, implying that services should be actively identifying and addressing unmet needs. Assessing and then meeting health and social care needs beneficially focuses attention on a wide range of aspects of the patient's life, and promotes holistic care planning.

Section 8.6.3 explored the implications of this new type of service. A preliminary testable model of the relationship between mental health interventions and quality of life was proposed in Section 8.6.3.1. The characteristics of a new type of mental health service based on this model were outlined in Section 8.6.3.2. The effectiveness of the new service could be explicitly compared with standard services using a randomised controlled trial, with patients randomly allocated to one of the two services. A service in which care was driven by the patient's assessment of their own needs would have a different philosophy of care, and so attention would need to be given in such a design to balancing potential confounding factors, such as staff preference (as per Section 8.5.5), the length of time the team has been in operation, the skill levels in team members, and so forth.

9.4 Policy implications

This study has two implications for mental health policy, relating to the policy on providing care based on need and on routine outcome assessment.

9.4.1 Implication 8: Providing care on the basis of need

The National Health Service and Community Care Act (1990) requires that care be provided on the basis of need, and evaluated in terms of its impact on quality of life. The linking of level of care with level of need is also made in the Mental Health National Service Framework (1999). This policy is supported by the FOCUS RCT, which did find evidence that reductions in a patient's self-rated unmet needs led to a better self-rated quality of life. The broader policy implication of the model described in

Figure 8.7 is that there may be a need in the future to extend the remit – and hence the resourcing – of mental health services, if substantial change in quality of life is to become a reality.

9.4.2 Implication 9: Routine outcome assessment

The drive towards routine and widespread use of clinical outcomes continues (Department of Health, 2001). Routine outcome assessment is the subject of a recent academic debate following a pessimistic appraisal of the usefulness of routine outcome measurement for planners of mental health services (Bilsker & Goldner, 2002). In response, Trauer (2003) argued that respondent bias can be reduced through triangulation of multiple responses, and Callaly, Coombs and Berk (2003) identified the benefits of using outcome measures routinely within the clinical care-planning process. Underpinning this debate is the need for clarity about the purpose of routine outcome measurement.

One purpose would be to aggregate data to allow the local benchmarking of teams and services, and so to inform efficient resource allocation. A second purpose would be to facilitate national comparisons, both between service models and between different populations, and hence to inform service development priorities. The FOCUS Study does not directly inform either of these uses. However, the level of implementation for a reasonably intense intervention in the FOCUS RCT, as described in Section 8.3, does contrast with the difficulties in obtaining high-quality activity data (Hansell, Bottle, Shurlock et al, 2001). This suggests that it may be easier to get adequate data quality where the provider of the information is also the user of the information. Where the information provider does not use the information (as with activity data), adequate data quality may in practice be impossible to achieve.

The third possible purpose of encouraging routine use of outcomes may be to inform the care of individual patients. If this is the policy goal, then the FOCUS RCT will have high relevance – it is the first randomised controlled trial to involve staff and patients in routine and repeated collection and feedback of outcome information. The results indicate that achieving health gains in adult mental health services will require a different design, and probably more resources, than the FOCUS RCT. This result is consistent with other emerging evidence investigating routine use of health-related quality of life, needs and symptom measures (Ashaye et al, 2003; Gilbody et al, 2001;

Gilbody et al, 2002b). Currently available evidence suggests that national implementation of routine collection of outcome data with the intention of improving the care of individual patients would be premature.

References

Acheson D (1998) *Inequalities in Health*, London: HMSO.

Altman D (1991) *Practical Statistics for Medical Research*, London: Chapman & Hall.

Altman DA, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gøtzsche PC, Lang T for the CONSORT Group (2001) *The Revised CONSORT Statement for Reporting Randomized Trials: Explanation and Elaboration*, Annals of Inernal Medicine, **134**, 663-694.

American Psychiatric Association (2000) *Handbook of Psychiatric Measures*, Washington DC: American Psychiatric Association.

American Psychiatric Association Task Force on Quality Indicators (2002) *Quality indicators*, Washington DC: American Psychiatric Association.

Anderson JR (1983) *The Architecture of Cognition*, Harvard: Harvard University Press.

Andresen R, Caputi P, Oades LG (2000) *Interrater reliability of the Camberwell Assessment of Need Short Appraisal Schedule*, Australian and New Zealand Journal of Psychiatry, **34**, 856-861.

Andrews G (1999) *Efficacy, effectiveness and efficiency in mental health services delivery*, Australian and New Zealand Journal of Psychiatry, **33**, 316-322.

Andrews G, Henderson S (Eds.) (2000) *Unmet need in psychiatry*, Cambridge: Cambridge University Press.

Andrews G, Peters L, Teesson M (1994) *Measurement of consumer outcome in mental health: A report to the National Mental Health Information Strategy Committee*, Sydney: Clinical Research Unit for Anxiety Disorders.

Ashaye OA, Livingston G, Orrell MW (2003) *Does standardized needs assessment improve the outcome of psychiatric day hospital care for older people? A randomized controlled trial*, Aging and Mental Health, **7**, 195-199.

Attkisson C, Cook J, Karno M, Lehman A, McGlashan TH, Meltzer HY, O'Connor M, Richardson D, Rosenblatt A, Wells K, Williams J, Hohmann AA (1992) *Clinical services research*, Schizophrenia Bulletin, **18**, 561-626.

Baker F, Intagliata J (1982) *Quality of life in the evaluation of community support systems*, Evaluation and Program Planning, **5**, 69-79.

Baron RM, Kenny DA (1986) *The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations*, Journal of Personality and Social Psychology, **51**, 1173-1182.

Barr W (2000) *Characteristics of severely mentally ill patients in and out of contact with community mental health services*, Journal of Advanced Nursing, **31**, 1189-1198.

Barry MM, Zissi A (1997) *Quality of life as an outcome measure in evaluating mental health services: a review of the empirical evidence*, Social Psychiatry and Psychiatric Epidemiology, **32**, 38-47.

Bartlett CJ, Coles E (1998) *Psychological health and well-being: why and how should public health specialists measure it? Part 2: stress, subjective well-being and overall conclusions*, Journal of Public Health Medicine, **20**, 288-294.

Beauford JE, McNiel DE, Binder RL (1997) *Utility of the initial therapeutic alliance in evaluating psychiatric patients' risk of violence*, American Journal of Psychiatry, **154**, 1272-1276.

Bech P (1993) *Rating scales for psychopathology, health status and quality of life*, Berlin: Springer-Verlag.

Beck AT, Rush AJ, Shaw BF, Emery G (1979) *Cognitive therapy of depression*, New York: Guilford Press.

Beck CA, Joseph L, Belisle P, Pilote L (2002) *Predictors of quality of life 6 months and 1 year after acute myocardial infarction*, American Heart Journal, **142**, 271-279.

Becker M, Diamond R, Sainfort F (1993) *A new patient focused index for measuring quality of life in persons with severe and persistent mental illness*, Quality of Life Research, **2**, 239-251.

Bengtsson-Tops A, Hansson L (1999) *Clinical and social needs of schizophrenic outpatients living in the community: the relationship between needs and subjective quality of life*, Social Psychiatry and Psychiatric Epidemiology, **34**, 513-518.

Benjamin K, Perfetto E, Greene J (1995) *Public policy and the application of outcomes assessments: paradigms versus politics*, Medical Care, **33**, AS299-AS305, suppl.

Bergner M (1985) *Measurement of health status*, Medical Care, **23**, 696-704.

Bergner M, Rothman ML (1987) *Health status measures: an overview and guide for selection*, American Review of Public Health, **8**, 191-210.

Biggeri A, Bisoffi G, Rucci P, Ruggeri M, Tansella M (2001) *Graphical models for the multidimensional assessment of outcome*. In M Tansella, G Thornicroft (Eds.) "Mental Health Outcome Measures", 2nd edition, London: Gaskell, pp. 258-270.

Bilsker D, Goldner EM (2002) *Routine outcome measurement by mental health-care providers: is it worth doing?*, Lancet, **360**, 1689-1690.

Bindman J, Beck A, Glover G, Thornicroft G, Knapp M, Leese M, Szmukler G (1999) *Evaluating mental health policy in England. Care Programme Approach and supervision registers*, British Journal of Psychiatry, **175**, 327-330.

Bindman J, Glover G, Goldberg D, Chisholm D (2000) *Expenditure on mental health care by English health authorities: a potential cause of inequity*, British Journal of Psychiatry, **177**, 267-274.

Bindman J, Johnson S, Wright S, Szmukler G, Bebbington P, Kuipers E, Thornicroft G (1998) *Integration between primary and secondary services in the care of the severely mentally ill: patients and general practitioners' views*, British Journal of Psychiatry, **171**, 169-174.

Boardman AP, Hodgson RE, Lewis M, Allen K. (1999) *North Staffordshire Community Beds Study: longitudinal evaluation of psychiatric in-patient units attached to community mental health centres. I: Methods, outcome and patient satisfaction*, British Journal of Psychiatry, **175**, 70-78.

Bollen K (1989) *Structural equations with latent variables*, New York: Wiley.

Bolton D, Hill J (1996) *Mind, meaning and mental disorder*, Oxford: Oxford University Press.

Borge L, Martinsen EW, Ruud T, Watne O, Friis S (2000) *Quality of life, loneliness and social contacts among persons with long-term mental illness* [Norwegian], Tidsskrift for Den Norske Laegeforening, **120**, 52-55.

Bowling A (1991) *Measuring health: A review of quality of life measurement scales*, Milton Keynes: Open University Press.

Bowling, A (1997) *Research Methods in Health: investigating health and health services*, Milton Keynes: Open University Press.

Bracken P, Thomas P (2001) *Postpsychiatry: a new direction for mental health,* BMJ, **322**, 724-727.

Bradley C (1993) *Designing medical and educational intervention studies*, Diabetes Care, **16**, 509-518.

Brewin C, Wing J, Mangen S, Brugha T, MacCarthy B (1987) *Principles and practice of measuring needs in the long-term mentally ill: the MRC Needs for Care Assessment*, Psychological Medicine, **17**, 971-981.

Brooks R (2000) *The reliability and validity of the Health of the Nation Outcome Scales: validation in relation to patient derived measures*, Australian and New Zealand Journal of Psychiatry, **34**, 504-511.

Brown S, Inskip H, Barraclough B (2000) *Causes of the excess mortality of schizophrenia*, British Journal of Psychiatry, **177**, 212-217.

Buchanan DR (1992) *An uneasy alliance: combining qualitative and quantitative research methods*, Health Education Quarterly, Spring, 117-135.

Burnam MA (1996) *Measuring outcomes of care for substance use and mental disorders*, New Directions for Mental Health Services, **71**, 3-17.

Burns T, Priebe S (1996) *Mental health care systems and their characteristics: a proposal*, Acta Psychiatrica Scandinavica, **94**, 381-385.

Callaly T, Coombs T, Berk M (2003) *Routine outcome measurement by mental health-care providers*, Lancet, **361**, 1137-1138.

Campbell J (1996) *Towards collaborative mental health outcome systems*, New Directions for Mental Health Services, **71**, 69-78.

Campbell J (1998) *Consumerism, Outcomes and Satisfaction: A review of the literature*, accessed at www.madnation.org/citations/consumerism.htm.

Campbell M, Fitzpatrick R, Haines A, Kinmouth AL, Sandercock P, Spiegelhalter D, Tyrer P (2000) *Framework for design and evaluation of complex interventions to improve health*, BMJ, **321**, 694-696.

Chalmers I (2001) *Comparing like with like: some historical milestones in the evolution of methods to create unbiased comparison groups in therapeutic experiments*, International Journal of Epidemiology, **30**, 1156-1164.

Ciarlo J, Edwards D, Kiresuk T, Newman F, Brown T (1981) *The assessment of client/patient outcomes techniques for use in mental health programs*, Washington DC: National Institute of Mental Health.

Clifford P (1998) *M is for outcome: The CORE outcomes initiative*, Journal of Mental Health, **7**, 19-24.

Clifford P (1999) *The FACE Recording and Measurement System: A scientific approach to person-based information*, Bulletin of the Menninger Clinic, **63**, 305-331.

Cook D, Sackett D, Spritzer W (1995) *Methodologic guidelines for systematic reviews of randomized controlled trials in health care from the Potsdam consultation on meta-analysis*, Journal of Clinical Epidemiology, **48**, 167-171.

Cook J, Jonikas J (1996) *Outcomes of psychiatric rehabilitation service delivery*, New Directions for Mental Health Services, **71**, 33-47.

Corrigan PW, Buican B (1995) *The construct validity of subjective quality of life for the severely mentally ill*, Journal of Nervous and Mental Disease, **183**, 281-285.

Corrigan PW, Lickey SE, Campion J, Rashid F (2000) *Mental health team leadership and consumers satisfaction and quality of life*, Psychiatric Services, **51**, 781-785.

Coste J, Guillemin F, Pouchot J, Fermanian J (1997) *Methodological Approach to Shortening Composite Measurement Scales*, Journal of Clinical Epidemiology, **50**, 247-252.

Crawford MJ, Rutter D, Manley C, Weaver T, Bhui K, Fulop T, Tyrer P (2002) *Systematic review of involving patients in the planning and development of health care*, BMJ, **325**, 1263-1267.

Cummins R (1995) *On the trail of the gold standard for subjective well-being*, Social Indicators Research, **35**, 179-200.

Deeks J, Khan KS, Popay J, Nixon J, Kleijnen J (2000) *Stage III Conducting the review: Phase 7 Data synthesis*. In "Undertaking systematic reviews of research on effectiveness", NHS Centre for Reviews and Dissemination Report No. 4, 2nd edition, University of York, pp. 56-74.

Department of Health (1998a) *Modernising Mental Health Services*, London: Department of Health.

Department of Health (1998b) *Our Healthier Nation*, London: The Stationery Office.

Department of Health (1999) *Mental Health National Service Framework*, London: The Stationery Office.

Department of Health (2001) *Mental Health Information Strategy*, London: The Stationery Office.

Dickerson E (1997) *Assessing clinical outcomes: the community functioning of persons with serious mental illness*, Psychiatric Services, **48**, 897-902.

Dickey B, Azeni H (1992) *Data watch: Impact of managed care on mental health services*, Health Affairs, **11**, 197-204.

Dickey B, Sederer LI (Eds.) (2001) *Improving mental health care*, Washington DC: American Psychiatric Association.

Doyle M, Flanagan S, Browne S, Clarke M, Lydon D, Larkin E, O'Callaghan C (1999) *Subjective and external assessments of quality of life in schizophrenia: relationship to insight*, Acta Psychiatrica Scandinavica, **96**, 466-472.

Drury V, Birchwood M, Cochrane R, Macmillan F (1996) *Cognitive therapy and recovery from acute psychosis: a controlled trial. I. Impact on psychotic symptoms*, British Journal of Psychiatry, **169**, 593-601.

Dunn G (2001) *Statistical methods for measuring outcomes.* In G Thornicroft, M Tansella (Eds.) "Mental Health Outcome Measures", 2nd edition, London: Gaskell, pp. 5-18.

Edwards D (2000) *Introduction to Graphical Modelling*, 2nd edition, New York: Springer Verlag.

Eisen SV, Dill DL, Grob MC (1994) *Reliability and validity of a brief patient-report instrument for psychiatric outcome evaluation*, Hospital and Community Psychiatry, **45**, 242-247.

Ellwood P (1988) *Outcomes Management – A Technology of Patient Experience*, New England Journal of Medicine, **318**, 1549-1556.

Endicott J, Nee J, Harrison W, Blumenthal R (1993) *Quality of Life Enjoyment and Satisfaction Questionnaire: a new measure*, Psychopharmacology Bulletin, **29**, 321-326.

Endicott J, Spitzer RL (1976) *The Global Assessment Scale: a procedure for measuring overall severity of psychiatric disturbance*, Archives of General Psychiatry, **33**, 766-771.

Espellargues M, Valderas JM, Alonso J (2000) *Provision of Feedback on Perceived Health Status to Health Care Professionals*, Medical Care, **38**, 175-186.

Essink-Bot M-L, Krabbe PFM, Bonsel GJ, Aaronson NK (1997) *An empirical comparison of four generic health status measures*, Medical Care, **35**, 522-537.

Evans S, Greenhalgh J, Connelly J (2000) *Selecting a mental health needs assessment scale: guidance on the critical appraisal of standardized measures*, Journal of Evaluation in Clinical Practice, **6**, 379-393.

Evans S, Huxley P (2002) *Subjective Quality of Life in the General Population*, International Review of Psychiatry, **14**, 203-211.

Everitt BS, Wessely S (2004) *Clinical trials in psychiatry*, Oxford: Oxford University Press.

Fenton WS, Blyler CR, Heinssen RK, (1997) *Determinants of medication compliance in schizophrenia*, Schizophrenia Bulletin, **23**, 637-651.

Fitzpatrick R, Davey C, Buxton M, Jones D (1998) *Evaluating patient-based outcome measures for use in clinical trials*, Health Technology Assessment, **2(14)**.

Foldemo A, Bogren L (2002) *Need assessment and quality of life in outpatients with schizophrenia: a 5-year follow-up study*, Scandinavian Journal of Caring Sciences, **16**, 393-398.

Garety P, Fowler D, Kuipers E, Freeman D, Dunn G, Bebbington P, Hadley C, Jones S (1997) *London-East Anglia randomised controlled trial of cognitive-behavioural therapy for psychosis. II: Predictors of outcome*, British Journal of Psychiatry, **171**, 420-426.

Gilbody SM, House AO, Sheldon TA (2001) *Routinely administered questionnaires for depression and anxiety: systematic review*, BMJ, **322**, 406-409.

Gilbody SM, House AO, Sheldon TA (2002a) *Psychiatrists in the UK do not use outcome measures*, British Journal of Psychiatry, **180**, 101-103.

Gilbody SM, House AO, Sheldon TA (2002b) *Routine administration of Health Related Quality of Life (HRQoL) and needs assessment instruments to improve psychological outcome – a systematic review*, Psychological Medicine, **322**, 406-409.

Gilbody S, Whitty P (2002) *Improving the delivery and organisation of mental health services: beyond the conventional randomised controlled trial*, British Journal of Psychiatry, **180**, 13-18.

Glover G (1997) *The development of a new minimum data set for specialist mental health*, Health Trends, **29**, 48-51.

Glover GR, Robin E, Emami J, Arabscheibani GR (1998) *A needs index for mental health care*, Social Psychiatry and Psychiatric Epidemiology, **33**, 89-96.

Gold JM, Goldberg RW, McNary SW, Dixon LB, Lehman AF (2002) *Cognitive correlates of job tenure among patients with severe mental illness*, American Journal of Psychiatry, **159**, 1395-1402.

Grahn B, Ekdahl C, Borgquist L (2002) *Motivation as a predictor of changes in quality of life and working ability in multidisciplinary rehabilitation*, Disability Rehabilitation, **22**, 639-654.

Green L, Eriksen M (1988) *Behavioural determinants of preventive practices by physicians*, American Journal of Preventative Medicine, **4**, suppl. 1, 101-107.

Green RS, Gracely E (1987) *Selecting a rating scale for evaluating services to the chronically mentally ill*, Community Mental Health Journal, **23**, 91-102.

Grol R, Grimshaw J (2003) *From best evidence to best practice: effective implementation of change in patients' care*, Lancet, **362**, 1225-1230.

Guadagnoli E, Ward P (1998) *Patient participation in decision-making*, Social Science and Medicine, **47**, 329-339.

Hansell A, Bottle A, Shurlock L, Aylin P (2001) *Accessing and using hospital activity data*, Journal of Public Health Medicine, **23**, 51-56.

Hansson L, Sandlund M, Bengtsson-Tops A, Bjarnason O, Karlsson H, Mackeprang T, Merinder L, Nilsson L, Sorgaard K, Vinding H, Middelboe T (2003) *The relationship of needs and quality of life in persons with schizophrenia living in the community. A Nordic multi-center study*, Nordic Journal of Psychiatry, **57**, 5-11.

Hansson L, Vinding HR, Mackeprang T, Sourander A, Werdelin G, Bengtsson-Tops A, Bjarnason O, Dybbro J, Nilsson L, Sandlund M, Sorgaard K, Middelboe T (2001) *Comparison of key worker and patient assessment of needs in schizophrenic patients living in the community: a Nordic multicentre study*, Acta Psychiatrica Scandinavica, **103**, 45-51.

Hargreaves W, Shumway M (1989) *Effectiveness of mental health services for the severely mentally ill.* In CA Taube, D Mechanic, AA Hohmann (Eds.) "The future of mental health services research", Washington DC: US Government Printing Office, pp. 253-284.

Harrison G, Eaton W (1999) *From research world to real world*, Current Opinion in Psychiatry, **12**, 187-189.

Hassiotis A, Ukoumunne OC, Byford S, Tyrer P, Harvey K, Piachaud J, Gilvarry K, Fraser J (2001) *Intellectual functioning and outcome of patients with severe psychotic illness randomised to intensive case management*, British Journal of Psychiatry, **178**, 166-171.

Henkel H, Schmitz M, Berghofer G, Lang A, Kager A, Steiner E, Schmidl F, Rudas S (2000) *Quality of life of the mentally ill* [German], Wiener Medizinische Wochenschrift, **150**, 32-36.

Hoffmann K, Priebe S, Isermann M, Kaiser W (1997) *Quality of life, needs and treatment evaluation of long-term hospitalized patients. Part II of the Berlin Deinstitutionalization Study* [German], Psychiatrische Praxis, **24**, 221-226.

Holloway F, Carson J (1999) *Subjective quality of life, psychopathology, satisfaction with care and insight: an exploratory study*, International Journal of Social Psychiatry, **45**, 259-267.

Horwitz AV, Uttaro T (1998) *Age and mental health services*, Community Mental Health Journal, **34**, 275-287.

Howard KI, Moras K, Brill P, Martinovich Z, Lutz W (1996) *Efficacy, Effectiveness and Patient Progress*, American Psychologist, **51**, 1059-1064.

Howe K, Eisenhart M (1990) *Standards for qualitative (and quantitative) research: a prolegomenon*, Educational Researcher, **19**, 2-9.

Huxley P (1998) *Outcomes management in mental health: A brief review*, Journal of Mental Health, **7**, 273-283.

Iles V, Sutherland K (2001) *Organisational Change*, London: National Co-ordinating Centre for Service Development and Organisation.

Independent Inquiry into the death of David Bennett (2003) Cambridge: Norfolk, Suffolk and Cambridgeshire Strategic Health Authority.

IsHak WWI, Burt T, Sederer LI (Eds.) (2002) *Outcome Measurement in Psychiatry*, Washington DC: American Psychiatric Association.

Issakidis C, Teesson M (1999) *Measurement of need for care: a trial of the Camberwell Assessment of Need and the Health of the Nation Outcome Scales*, Australian and New Zealand Journal of Psychiatry, **33**, 754-759.

Jausovec N (1987) *Differences in the solution process of ill- and well-defined problems in literature*, Anthropos, **1**, 237-249.

Jennings BM, Staggers N, Brosch LR (1999) *A classification scheme for outcome indicators*, Image: Journal of Nursing Scholarship, **31**, 381-388.

Kahney H (1986) *Problem solving: a cognitive approach*, Milton Keynes: Open University Press.

Kane RA, Kane RL, Arnold S (1985) *Measuring social functioning in mental health studies: concepts and instruments*, Rockville: NIMH.

Katschnig H (1997) *How useful is the concept of quality of life in psychiatry?*, Current Opinion in Psychiatry, **10**, 337-345.

Keijsers GP, Schaap CP, Hoogduin CA (2000) *The impact of interpersonal patient and therapist behavior on outcome in cognitive-behavioural therapy. A review of empirical studies*, Behavior Modification, **24**, 264-297.

Khan KS, ter Riet G, Popay J, Nixon J, Kleijnen J (2000) *Phase 5: Study quality assessment*. In "Undertaking systematic reviews of research on effectiveness", NHS Centre for Reviews and Dissemination Report No. 4, 2nd edition, University of York, pp. 37-50.

Koivumaa-Honkanen HT, Viinamaki H, Honkanen R, Tanskanen A, Antikainen R, Niskanen L, Jaaskelainen J, Lehtonen J (1996) *Correlates of life satisfaction among psychiatric patients*, Acta Psychiatrica Scandinavica, **94**, 372-378.

Kolb DA (1984) *Experiential Learning: Experience as the source of learning and development*, New Jersey: Prentice-Hall.

Krupnick JL, Sotsky SM, Simmens S, Moyer J, Elkin I, Watkins J, Pilkonis PA (1996) *The role of therapeutic alliance in psychotherapy and pharmacotherapy outcome:*

findings in the National Institute of Mental Health Treatment of Depression Collaborative Research Program, Journal of Consulting and Clinical Psychology, **64**, 532-539.

Lam JA, Rosenheck RA (2000) *Correlates of improvement in quality of life among homeless persons with serious mental illness*, Psychiatric Services, **51**, 116-118.

Lambert MJ, Hansen NB, Umphress V, Lunnen K, Okiishi J, Burlingame G, Huefner JC, Reisinger CW (1996) *Administration and scoring manual for the Outcome Questionnaire (OC 45.2)*, Wilmington DE: American Professional Credentialing Services.

Lambert MJ, Whipple JL, Smart DA, Vermeersch DA, Nielson SL, Hawkins EJ (2001) *The effects of providing therapists with feedback on patient progress during psychotherapy: are outcomes enhanced?*, Psychotherapy Research, **11**, 49-68.

Lasalvia A, Ruggeri M, Mazzi MA, Dall'Agnola RB (2000) *The perception of needs for care in staff and patients in community-based mental health services. The South Verona Outcome Project 3*, Acta Psychiatrica Scandinavica, **102**, 366-375.

Laugharne RA (1999) *Evidence based medicine, user involvement and the postmodern paradigm*, Psychiatric Bulletin, **23**, 641-643.

Leese M, Johnson S, Slade M, Parkman S, Kelly F, Phelan M, Thornicroft G (1998) *The User Perspective on Needs and Satisfaction with Mental Health Services: the PRiSM Psychosis Study (8)*, British Journal of Psychiatry, **173**, 409-415.

Lefebvre J, Cyr M, Lesage A, Fournier L, Toupin J (2000) *Unmet needs in the community: can existing services meet them?*, Acta Psychiatrica Scandinavica, **102**, 65-70.

Lehman AF (1983) *The effects of psychiatric symptoms on quality of life assessment among the chronic mentally ill*, Evaluation and Program Planning, **6**, 143-151.

Lehman AF (1988) *A quality of life interview for the chronically mentally ill*, Evaluation and Program Planning, **11**, 51-62.

Lehman AF, Kernan E, DeForge BR, Dixon L, (1995) *Effects of homelessness on the quality of life of persons with severe mental illness*, Psychiatric Services, **46**, 922-926.

Lelliott P (2000) *What do people want from specialist mental health services and can this be routinely measured in routine service settings?*, Behavioural and Cognitive Psychotherapy, **28**, 361-368.

Lelliott P, Beevor A, Hogman J, Hogman G, Hyslop J, Lathlean J, Ward M (2001) *Carers' and Users' Expectations of Services – User version (CUES-U): a new*

instrument to measure the experience of users of mental health services, British Journal of Psychiatry, **179**, 67-72.

Lockwood A, Marshall M (1999) *Can a standardized needs assessment be used to improve the care of people with severe mental disorders? A pilot study of 'needs feedback'*, Journal of Advanced Nursing, **30**, 1408-1415.

Long A, Dixon P (1996) *Monitoring outcomes in routine practice: defining appropriate measurement criteria*, Journal of Evaluation in Clinical Practice, **2**, 71-78.

Lyons JS, Howard KI, O'Mahoney MT, Lish JD (1997) *The measurement and management of clinical outcomes in mental health*, New York: John Wiley & Sons.

Macpherson R, Jerrom B, Lott G, Ryce M (1999) *The outcome of goal setting in a mental health rehabilitation service. A model for evaluating clinical effectiveness*, Journal of Mental Health, **8**, 95-102.

Margison FR, Barkham M, Evans C, McGrath G, Clark JM, Audin K, Connell J (2000) *Measurement and psychotherapy. Evidence-based practice and practice-based evidence*, British Journal of Psychiatry, **177**, 123-130.

Marks I (1998) *Overcoming obstacles to routine outcome measurement*, British Journal of Psychiatry, **173**, 281-286.

Marshall M, Hogg L, Gath DH, Lockwood A (1995) *The Cardinal Needs Schedule: a modified version of the MRC Needs for Care Schedule*, Psychological Medicine, **25**, 605-617.

Martin DJ, Garske JP, Davis MK (2000) *Relation of the therapeutic alliance with outcome and other variables: a meta-analytic review*, Journal of Consulting and Clinical Psychology, **68**, 438-450.

McCabe R, Roeder-Wanner U-U, Hoffmann K, Priebe S (1999) *Therapeutic relationships and quality of life: Association of two subjective constructs in schizophrenia patients*, International Journal of Social Psychiatry, **45**, 276-283.

McColl E, Jacoby A, Thomas L, Soutter J, Bamford C, Steen N, Thomas R, Harvey E, Garratt A, Bond J (2001) *Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients*, Health Technology Assessment, **5**, 31.

McCrone P, Leese M, Thornicroft G, Schene AH, Knudsen HC, Vázquez-Barquero JL, Lasalvia A, Padfield S, White IR, Griffiths G and the EPSILON Study Group (2000) *Reliability of the Camberwell Assessment of Need – European Version. EPSILON Study 6*, British Journal of Psychiatry, **177**, suppl. 39, s34-s40.

McCrone P, Strathdee G (1994) *Needs not diagnosis: towards a more rational approach to community mental health resourcing in Britain*, International Journal of Social Psychiatry, **40**, 79-86.

McDowell I, Newell C (1987) *Measuring health: a guide to rating scales and questionnaires*, Oxford: Oxford University Press.

McGlynn E (1996) *Setting the context for measuring patient outcomes*, New Directions for Mental Health Services, **71**, 19-32.

McPheeters HL (1984) *Statewide mental health outcome evaluation: A perspective of two southern states*, Community Mental Health Journal, **20**, 44-55.

McPherson K, Britton AR, Wennberg JE (1997) *Are randomized controlled trials controlled? Patient preference and unblind trials*, Journal of the Royal Society of Medicine, **90**, 652-656.

Medical Research Council (1998) *MRC Guidelines for Good Clinical Practice in Clinical Trials*, London: MRC.

Medical Research Council (2000) *A framework for development and evaluations of RCTs for complex interventions to improve health*, London: MRC.

Meehl P, Golden RR (1982) *Taxonometric methods*. In PC Kendall, JN Butcher (Eds.) "Handbook of research methodology in clinical psychology", New York: Wiley, pp. 127-181.

Mercier C, Peladeau N, Tempier R (1998) *Age, gender and quality of life*, Community Mental Health Journal, **34**, 487-500.

Michalos AC, Zumbo BD (2001) *Ethnicity, modern prejudice and the quality of life*, Social Indicators Research, **53**, 189-222.

Moher D, Cook D, Eastwood S, Olkin I, Rennie D, Stroup F (1999) *Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement*, Lancet, **354**, 1896-1900.

Munro JC, Russell AJ, Murray RM, Kerwin RW, Jones PB (2002) *IQ in childhood psychiatric attendees predicts outcome of later schizophrenia at 21 year follow-up*, Acta Psychiatrica Scandinavica, **106**, 139-142.

Nazroo JY (1999) *Rethinking the relationship between ethnicity and mental health: The British Fourth National Survey of Ethnic Minorities*, Social Psychiatry, **33**, 145-148.

Nelson EC, Landgraf JM, Hays RD, Wasson JH, Kirk JW (1990) *The functional status of patients. How can it be measured in physicians' offices?*, Medical Care, **28**, 1111-1126.

Nelson HE (1982) *National Adult reading Test (NART): Test Manual*, Windsor: NFER-Nelson.

Nutley S, Percy-Smith J, Solesbury W (2003) *Models of research impact*, London: Learning and Skills Research Centre.

Oakley A (2000) *Experiments in Knowing: Gender and Method in Social Sciences*, Cambridge: Polity Press.

Oliver J (1991) *The social care directive: Development of a quality of life profile for use in community services for the severely mentally ill*, Social Work and Social Sciences Review, **3**, 5-45.

Oliver JPJ, Huxley PJ, Priebe S, Kaiser W (1997) *Measuring the quality of life of severely mentally ill people using the Lancashire Quality of Life Profile*, Social Psychiatry and Psychiatric Epidemiology, **32**, 76-83.

Overall JE, Gorham DR (1988) *The Brief Psychiatric Rating Scale (BPRS): Recent Developments in Ascertainment and Scaling*, Psychopharmacology Bulletin, **24**, 97-99.

Pawson R, Tilley N (1997) *Realistic Evaluation*, London: Sage.

Perese EF (1997) *Unmet needs of persons with chronic mental illnesses: relationship to their adaptation to community living*, Issues in Mental Health Nursing, **18**, 19-34.

Phelan M, Slade M, Thornicroft G, Dunn G, Holloway F, Wykes T, Strathdee G, Loftus L, McCrone P, Hayward P (1995) *The Camberwell Assessment of Need: the validity and reliability of an instrument to assess the needs of people with severe mental illness*, British Journal of Psychiatry, **167**, 589-595.

Preston NJ (2000) *The Health of the Nation Outcome Scales: validating factorial structure and invariance across two health services*, Australian and New Zealand Journal of Psychiatry, **34**, 512-519.

Priebe S, Gruyters T (1993) *The role of the helping alliance in psychiatric community care: A prospective study*, Journal of Nervous and Mental Disease, **181**, 552-557.

Priebe S, Huxley P, Knight S, Evans S (1999) *Application and results of the Manchester Short Assessment of quality of life*, International Journal of Social Psychiatry, **45**, 7-12.

Priebe S, McCabe R, Bullenkamp J, Hansson L, Roessler W, Torres-Gonzales F, Wiersma D (2002) *The impact of routine outcome measurement on treatment processes in community mental health care: Approach and methods of the MECCA study*, Epidemiologia e Psichiatria Sociale, **11**, 198-205.

Prior P (1999) *Gender and mental health*, New York: New York University Press.

Reynolds T, Thornicroft G, Abas M, Woods B, Hoe J, Leese M, Orrell M (2000) *Camberwell Assessment of Need for the Elderly (CANE); development, validity, and reliability*, British Journal of Psychiatry, **176**, 444-452.

Ritsner M, Modai I, Endicott J, Rivkin O, Nechamkin Y, Barak P, Goldin V, Ponizovsky A (2000) *Differences in quality of life domains and psychopathologic and psychosocial factors in psychiatric patients*, Journal of Clinical Psychiatry, **61**, 880-889.

Rock D, Preston N (2001) *HoNOS: is there any point in training clinicians?*, Journal of Psychiatric and Mental Health Nursing, **8**, 405-409.

Roder-Wanner UU, Priebe S (1998) *Objective and subjective quality of life of first-admitted women and men with schizophrenia*, European Archives of Psychiatry and Clinical Neuroscience, **248**, 250-258.

Rohland BM, Langbehn DR, Rohrer JE (2000) *Relationship between service effectiveness and satisfaction among persons receiving Medicaid mental health services*, Psychiatric Services, **51,** 248-250.

Rosen A, Hadzi-Pavlovic D, Parker G (1989) *The life skills profile: a measure assessing function and disability in schizophrenia*, Schizophrenia Bulletin, **15**, 325-337.

Rosenblatt A, Attkisson C (1993) *Assessing outcomes for sufferers of severe mental disorder: a conceptual framework and review*, Evaluation and Program Planning, **16**, 347-363.

Rosenheck RA, Dausey DJ, Frisman L, Kasprow W (2000) *Outcomes after initial receipt of social security benefits among homeless veterans with mental illness*, Psychiatric Services, **51**, 1549-1554.

Roth A, Fonagy P (1997) *What works for whom?*, New York: Guilford Press.

Ruggeri M, Biggeri A, Rucci P, Tansella M (1998) *Multivariate analysis of outcome of mental health care using graphical chain models. The South-Verona Outcome Project 1*, Psychological Medicine, **28**, 1421-1431.

Ruggeri M, Bisoffi G, Fontecedro L, Warner R (2001) Subjective and objective dimensions of quality of life in psychiatric patients: a factor analytical approach. The South Verona Outcome Project 4, British Journal of Psychiatry, **178**, 268-275.

Ruggeri M, Dall'Agnola R (1993) The development and use of the Verona Expectations for Care Scale (VECS) and the Verona Service Satisfaction Scale (VSSS) for measuring expectations and satisfaction with community-based psychiatric services in patients, relatives and professionals, Psychological Medicine **23**, 511-523.

Ruggeri M, Tansella M (1995) *Evaluating outcome in mental health care*, Current Opinion in Psychiatry, **8**, 116-121.

Sainfort F, Becker M, Diamond R (1996) *Judgments of quality of life of individuals with severe mental disorders: patient self-report versus provider perspectives*, American Journal of Psychiatry, **153**, 497-502.

Salvador-Carulla L (2001) *Assessment Instruments in Psychiatry: Description and Psychometric Properties*. In M Tansella, G Thornicroft G (Eds.) "Mental Health Outcome Measures", 2nd edition, London: Gaskell, pp. 228-257.

Schafer M (1999) *Nomothetic and idiographic methodology in psychiatry – a historical-philosophical analysis*, Medicine, Health Care and Philosophy, **2**, 265-274.

Schalock RL (2000) *Outcome-Based Evaluation*, New York: Kluwer Academic.

Schlosser B (1996) *New perspectives on outcomes assessment: the philosophy and application of the subjective health process model*, Psychotherapy, **33**, 284-304.

Schraw G, Dunkle ME, Bendixen LD (1995) *Cognitive processes in well-defined and ill-defined problem solving*, Applied Cognitive Psychology, **9**, 523-538.

Schulz KF, Chalmers I, Hayes RJ, Altman DG (1995) Empirical evidence of bias: Dimensions of methodological quality associated with estimates of treatment effects in controlled trials, Journal of the American Medical Association, **273**, 408-412.

Schyns P (1998) Crossnational differences in happiness: Economic and cultural factors explored, Social Indicators Research, **43**, 3-26.

Seale S, Silverman D (1997) *Ensuring rigour in qualitative research*, European Journal of Public Health, **7**, 379-384.

Sederer LI, Hermann MD, Dickey B (1995) *The imperative of outcome assessment in psychiatry*, American Journal of Medical Quality, **10**, 127-132.

Shadish WR, Matt GE, Navarro AM, Siegle G, Crits-Christoph P, Hazelrigg MD, Jorm AF, Lyons LC, Nietzel MT, Prout HT, Robinson L, Smith ML, Svartberg M, Weiss B (1997) *Evidence that therapy works in clinically representative conditions*, Journal of Consulting and Clinical Psychology, **65**, 355-365.

Shern DL, Flynn LM (1996) *Outcomes roundtable: developing, evaluating and disseminating outcomes monitoring*, Behavioural Healthcare Tomorrow, **5**, 25-30.

Simpson A, Miller C, Bowers L (2003) *Case management models and the care programme approach: How to make the CPA effective and credible*, Journal of Psychiatric and Mental Health Nursing, **10**, 472-483.

Skantze K, Malm U, Dencker SJ, May PR, Corrigan P (1992) *Comparison of quality of life with standard of living in schizophrenia out-patients*, British Journal of Psychiatry, **161**, 797-801.

Slade M (1994) *Needs Assessment*, British Journal of Psychiatry, **165**, 293-296.

Slade M (2002) *What outcomes to measure in routine mental health services, and how to assess them – a systematic review*, Australian and New Zealand Journal of Psychiatry, **36**, 743-753.

Slade M, Beck A, Bindman J, Thornicroft G, Wright S (1999a) *Routine clinical outcome measures for severely mentally ill patients: CANSAS and HoNOS*, British Journal of Psychiatry, **174**, 404-408.

Slade M, Leese M, Taylor R, Thornicroft G (1999b) *The association between needs and quality of life in an epidemiologically representative sample of people with psychosis*, Acta Psychiatrica Scandinavica, **100**, 149-157.

Slade M, Loftus L, Phelan M, Thornicroft G Wykes T (1999c) *The Camberwell Assessment of Need*, London: Gaskell.

Slade M, Phelan M, Thornicroft G (1998) *A comparison of needs assessed by staff and an epidemiologically representative sample of patients with psychosis*, Psychological Medicine, **28**, 543-550.

Slade M, Phelan M, Thornicroft G, Parkman S (1996) *The Camberwell Assessment of Need (CAN): comparison of assessments by staff and patients of the needs of the severely mentally ill*, Social Psychiatry and Psychiatric Epidemiology, **31**, 109-113.

Slade M, Powell R, Rosen A, Strathdee G (2000) *Threshold Assessment Grid (TAG): the development of a valid and brief scale to assess the severity of mental illness*, Social Psychiatry and Psychiatric Epidemiology, **35**, 78-85.

Slade M, Thornicroft G, Glover G (1999d) *The feasibility of routine outcome measures in mental health*, Social Psychiatry and Psychiatric Epidemiology, **34**, 243-249.

Smith GR, Manderscheid R, Flynn L, Steinwachs D (1997) *Principles for assessment of patient outcomes in mental health care*, Psychiatric Services, **48**, 1033-1036.

Smith GR, Rost K, Fischer E, Burnam M, Burns B (1997) *Assessing the effectiveness of mental health care in routine clinical practice*, Evaluation and the helping professions, **20**, 65-80.

Smith KW, Avis NE, Assmann SF (1999) *Distinguishing between quality of life and health status in quality of life research: a meta-analysis*, Quality of Life Research, **8**, 447-459.

Spilker B (1990) *Quality of life assessments in clinical trials*, New York: Raven Press.

Spitzer RL, Kroenke K, Linzer M, Hahn SR, Williams JB, deGruy FV 3rd, Brody D, Davies M (1995) *Health-related quality of life in primary care patients with mental disorders. Results from the PRIME-MD 1000 Study*, Journal of the American Medical Association, **274**, 1511-1517.

Stedman T, Yellowlees P, Drake S, Chant D, Clarke R, Chapple B (2000) *The perceived utility of measures of consumer outcomes proposed for routine use in Australian mental health services*, Australian and New Zealand Journal of Psychiatry, **34**, 842-849.

Stedman T, Yellowlees P, Mellsop G, Clarke R, Drake S (1997) *Measuring consumer outcomes in mental health*, Canberra, ACT: Department of Health and Family Services.

Sterling T, Rosenbaum W, Weinkam J (1995) *Publication decisions revisited: the effect of the outcome of statistical tests on the decision to publish and vice versa*, American Statistician, **49**, 108-112.

Stewart AL, Hays RD, Ware JE (1988) *The MOS Short-Form General Health Survey: reliability and validity in a patient population*, Medical Care, **26**, 724-732.

Svensson B, Hansson L (1999) *Therapeutic alliance in cognitive therapy for schizophrenic and other long-term mentally ill patients: development and relationship to outcome in an in-patient treatment programme*, Acta Psychiatrica Scandinavica, **99**, 281-287.

Sytema S, Burgess P (1999) *Continuity of care and readmission in two service systems: a comparative Victoria and Groningen case-register study*, Acta Psychiatrica Scandinavica, **100**, 212-219.

Ten Have P (1999) *Doing conversation analysis: a practical guide*, London: Sage.

Thompson C (1989) *The instruments of psychiatric research*, Chichester: John Wiley & Sons.

Thornicroft G, Strathdee G, Phelan M, Holloway F, Wykes T, Dunn G, McCrone P, Leese M, Johnson S, Szmukler G (1998) *PRiSM Psychosis Study: Rationale and design*, British Journal of Psychiatry, **173**, 363-370.

Thornicroft G, Tansella M (1999) *The Mental Health Matrix*, Cambridge: Cambridge University Press.

Thornicroft G, Tansella M (Eds.) (2000) *Mental Health Outcome Measures*, 2nd edition, London: Gaskell.

Torgerson DJ, Sibbald B (1998) *Understanding controlled trials. What is a patient preference trial?*, BMJ, **316**, 360.

Trauer T (1998) *Issues in the assessment of outcome in mental health*, Australian and New Zealand Journal of Psychiatry, **32**, 337-343.

Trauer T (2003) *Routine outcome measurement by mental health-care providers*, Lancet, **361**, 1137.

Trauer T, Callaly T, Hantz P, Little J, Shields R, Smith J (1999) *Health of the Nation Outcome Scales. Results of the Victorian field trial*, British Journal of Psychiatry, **174**, 380-388.

UK700 Group (1999) *Predictors of quality of life in people with severe mental illness. Study methodology with baseline analysis in the UK700 trial*, British Journal of Psychiatry, **175**, 426-432.

UK Clearing House on Health Outcomes (1996) *Definitions of outcomes*, accessed at www.leeds.ac.uk/nuffield/infoservices/UKCH/define.html.

van Dongen CJ (1996) *Quality of life and self-esteem in working and nonworking persons with mental illness*, Community Mental Health Journal, **32**, 535-548.

van Minnen A, Hoogduin CA, Broekman TG (1997) *Hospital v. outreach treatment of patients with mental retardation and psychiatric disorders: a controlled study*, Acta Psychiatrica Scandinavica, **95**, 515-552.

van Os J, Altamura AC, Bobes J, Owens DC, Gerlach J, Hellewell JSE, Kasper S, Naber D, Robert P (2004) *Evaluation of the Two-Way Communication Checklist as a clinical intervention*, British Journal of Psychiatry, **184**, 79-83.

van Os J, Altamura AC, Bobes J, Owens DC, Gerlach J, Hellewell JSE, Kasper S, Naber D, Tarrier N, Robert P (2002) *2-com: an instrument to facilitate patient-professional communication in routine clinical practice*, Acta Psychiatrica Scandinavica, **106**, 446-452.

Varvin S (1991) *A retrospective follow-up investigation of a group of schizophrenic patients treated in a psychotherapeutic unit: The Kastanjebakken study*, Psychopathology, **24**, 336-344.

Veit CT, Ware JE (1983) *The structure of psychological distress and well-being in general populations*, Journal of Consulting and Clinical Psychology, **51**, 730-742.

Walter G, Cleary N, Rey J (1998) *Attitudes of mental health personnel towards rating outcome*, Journal of Quality in Clinical Practice, **18**, 109-115.

Ware J (1989) *Measuring health and functional status in mental health services research*. In C Taube, D Mechanic, A Hohmann (Eds.) "The future of mental health services research", Washington DC: US Government Printing Office, pp. 289-302.

Ware JE, Sherbourne CD (1992) *The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection*, Medical Care, **30,** 473-483.

Weiss B (1998) *Annotation: routine monitoring of the effectiveness of child psychotherapy*, Journal of Child Psychology and Psychiatry, **39**, 943-950.

Weiss M, Gaston L, Propst A, Wisebord S, Zicherman V (1997) *The role of the alliance in the pharmacologic treatment of depression*, Journal of Clinical Psychiatry, **58**, 196-204.

Weisz JR, Weiss B, Donenberg GR (1992) *The lab versus the clinic. Effects of child and adolescent psychotherapy*, American Psychologist, **47**, 1578-1585.

Wenger N, Mattson M, Furberg C, Elinson J (1984) *Assessment of Quality of Life in Clinical Trials of Cardiovascular Therapies*, New York: LeJacq Publishing.

White RE, Lyons J (1994) *Before report cards, we need road maps*, Modern Healthcare, **47**, 51-54.

Wiersma D, Nienhuis FJ, Giel R, de Jong A, Slooff CJ (1996) *Assessment of the need for care 15 years after onset of a Dutch cohort of patients with schizophrenia, and an international comparison*, Social Psychiatry and Psychiatric Epidemiology, **31**, 114-121.

Wiersma D, Nienhuis FJ, Giel R, Slooff CJ (1998) *Stability and change in needs of patients with schizophrenic disorders: a 15- and 17-year follow-up from first onset of psychosis, and a comparison between 'objective' and 'subjective' assessments of needs for care*, Social Psychiatry and Psychiatric Epidemiology, **33**, 49-56.

Wilkin D, Hallam L, Doggett M (1992) *Measures of need and outcome for primary health care*, Oxford: Oxford University Press.

Willitts M, Benzeval M, Stansfeld S (2004) *Partnership history and mental health over time*, Journal of Epidemiology and Community Health, **58**, 53-58.

Wing JK, Beevor AS, Curtis RH, Park SB, Hadden S, Burns A (1998) *Health of the Nation Outcome Scales (HoNOS) – Research and Development*, British Journal of Psychiatry, **172**, 11-18.

World Health Organization (1980) *International Classification of Impairments, Disabilities and Handicaps*, Geneva: World Health Organization.

World Health Organization (1998) *Disability Assessment Schedule (DAS-II)*, Geneva: World Health Organization.

Wright S, Gournay K, Glorney E, Thornicroft G (2000) *Dual diagnosis in the suburbs: prevalence, need, and in-patient service use*, Social Psychiatry and Psychiatric Epidemiology, **35**, 297-304.

Young A, Grusky O, Jordan D, Belin T (2000) *Routine outcome monitoring in a public mental health system: the impact of patients who leave care*, Psychiatric Services, **51**, 85-91.

Zuroff DC, Blatt SJ, Sotsky SM, Krupnick JL, Martin DL, Sanislow CA, Simmens S (2000) *Relation of therapeutic alliance and perfectionism to outcome in brief outpatient treatment of depression*, Journal of Consulting and Clinical Psychology, **68**, 114-224.

Appendix 1

This Appendix contains the forms used in the FOCUS RCT:

Patient information sheet and consent form 215

Staff information sheet and consent form 218

Postal questionnaire for patients,

comprising cover sheet, CANSAS-P, MANSA and HAS-P 220

Postal questionnaire for staff,

comprising cover sheet, CANSAS-S, TAG and HAS-S 225

Evaluation forms for patients, comprising Supplementary-P, NART

and BPRS 229

Evaluation forms for staff, comprising Supplementary-S and HoNOS 240

Sociodemographic form 245

Casenote form 248

Impact of Involvement form, comprising patient and staff versions 250

Adverse Event form 252

Admissions form 253

Demonstration feedback 254

Random number table 257

Institute of Psychiatry
at The Maudsley

Health Services Research Department
Section of Community Psychiatry (PRISM)

PO29
De Crespigny Park
Denmark Hill
London SE5 8AF
Telephone +44 (0)20 7848 0735
Fax +44 (0)20 7277 1462

KING'S College LONDON
University of London

Implementing routine outcome assessment in adult mental health services

PATIENT INFORMATION SHEET

Why is this study needed?

At present, people using mental health services do not always have their progress systematically and regularly monitored. Recent studies have found that there may be benefits for patients to chart their progress more closely. The goal of this study is to test whether monitoring progress more closely has positive benefits for patients.

What does this study involve?

At an initial meeting with a researcher, the purpose of the study will be explained, and you will be asked to sign a form to indicate that you give informed consent. If you agree, then you will be asked to fill in a short assessment form asking you about your needs, your quality of life and your relationship with your key-worker. This assessment form should take about 10 minutes, and is the form that will be sent to you each month. After you have filled in this form, you will be asked to complete some further assessments covering symptoms, service use and vocabulary. Your Casenotes will also be reviewed by a researcher.

After you have had this meeting, you will be randomly allocated to either the 'control' or the 'intervention' group. If you are put into the control group, then you will be contacted in 7 months time to re-assess your progress. If you are put into the intervention group, you will be sent the assessment form to complete every month. Every time you are asked to complete the assessment form, we will pay you £5 for your time. We will also ask your key-worker to complete a form every month, assessing their perceptions about your needs, the severity of your mental health problems, and their relationship with you. After three months and after six months you and your

key-worker will be sent identical feedback on the assessments you have both made over the previous three months. **This means that what you record on your forms will be seen by your key-worker.** We hope that this will stimulate discussion between you and your key-worker about what you agree or disagree about, the quality of your relationship, and whether the treatment and care you are receiving is helping you. After 7 months your progress will be re-assessed.

Confidentiality and consent

Participation in this study is entirely voluntary, and will not affect the care you receive from the mental health team. If you agree to take part, you may still withdraw from the study at any time, and your future care will not be affected.

All assessments will be confidential, apart from those which are fed back to you and your key-worker, as described above. You will not be identified in any presentation of the findings from this study. The only exception to this confidentiality will be if you disclose information which suggests a major risk of serious danger to any person, in which case your key-worker will be informed.

The project team

The project is being carried out by the FOCUS team, under the supervision of Dr Mike Slade (Chartered Clinical Psychologist). If you would like more information on this project, then please contact Dr Slade at the PRiSM team (Institute of Psychiatry) via 020 7848 5095.

If there is anything you do not understand in this form, then please ask. If you agree to take part in this study, then please sign the consent form.

Implementing routine outcome assessment in adult mental health services

PATIENT CONSENT FORM

The above study has been verbally described to me, and I have been given the written information sheet. I hereby give consent for my involvement in the above study. I understand that I can withdraw consent for involvement at any time, and that this will not impact on my care.

NAME .....

SIGNED .....

Witnessed by:

NAME .....

SIGNED .....

DATE .....

Institute of Psychiatry
at The Maudsley

Health Services Research Department
Section of Community Psychiatry (PRISM)

PO29
De Crespigny Park
Denmark Hill
London SE5 8AF
Telephone +44 (0)20 7848 0735
Fax +44 (0)20 7277 1462

KING'S College LONDON
University of London

Implementing routine outcome assessment in adult mental health services
STAFF INFORMATION SHEET

Why is this study needed?

At present, people using mental health services do not always have their progress systematically and regularly monitored. Recent studies have found that there may be benefits for patients to chart their progress more closely. The goal of this study is to test whether monitoring progress more closely has positive benefits for patients.

What does this study involve?

At an initial meeting with a researcher, the purpose of the study will be explained, and you will be asked to sign a form to indicate that you give informed consent. If you agree, then you will be asked to fill in a short assessment form asking you about their needs, the severity of their mental health problems and their relationship with you. This assessment form should take about 8 minutes, and is the form that will be sent to you each month. After you have filled in this form, you will be asked to complete further assessments covering sociodemographic and social disability information. The casenotes of the patient will be reviewed by the researcher.

After you have had this meeting, your patient will be randomly allocated to either the 'control' or the 'intervention' group. If they are put into the control group, then you will be contacted in 7 months time to re-assess their progress. If they are put into the intervention group, you will be sent the assessment form to complete every month. Every time you are asked to complete the assessment form, we will also ask your patient to complete a form assessing their perceptions about needs, quality of life and their relationship with you.

After three months and after six months you and your patient will be sent identical feedback on the assessments you have both made over the previous three months. **This means that what you record on your forms will be seen by your patient.** We hope that this will stimulate discussion between you and your patient about what you agree or disagree about, the quality of your relationship, and whether the treatment and care they are receiving is helping them. After 7 months their progress will be re-assessed.

Confidentiality and consent

Participation in this study is entirely voluntary. If you agree to take part, you may still withdraw from the study at any time. Non-participation will not affect your employment in any way. We will continue to ask you for information on your patient's progress, whether or not they choose to be involved in the study.

All assessments will be confidential, apart from those which are fed back to you and your patient, as described above. You and your patient will not be identified in any presentation of the findings from this study.

The project team

The project is being carried out by the FOCUS team, under the supervision of Dr Mike Slade (Chartered Clinical Psychologist). If you would like more information on this project, then please contact Dr Slade at the PRiSM team (Institute of Psychiatry) via 020 7848 5095.

If there is anything you do not understand in this form, then please ask. If you agree to take part in this study, then please sign the consent form.

Implementing routine outcome assessment in adult mental health services

STAFF CONSENT FORM

The above study has been verbally described to me, and I have been given the written information sheet. I hereby give consent for my involvement in the above study. I understand that I can withdraw consent for involvement at any time, and that this will not have any implications for my employment.

NAME .....

SIGNED .....

DATE .....

Witnessed by:

NAME .....

SIGNED .....

DATE .....

Feedback of Outcome to Users and Staff

Your name

Please fill in the questionnaire about…

Page 1. Your needs

Page 2. Your satisfaction with life

Page 3. Your relationship with your member of staff who is…

Please return the questionnaire in the FREEPOST envelope we have enclosed (no need for a stamp).

(Return address: FOCUS Study, Box P029, Institute of Psychiatry FREEPOST LON15630, London, SE5 8BR)

To discuss the study please contact Matthew or Lisa on 020 7848 5095.
However if you want to discuss any issues in confidence that have arisen from filling in the questionnaire you can phone:

- **MIND on 020 8668 2210 (for emotional help and advice) or**
- **Croydon Advocacy Service on 020 8665 9448**

"Croydon Advocacy Service can help you either by supporting you when you wish to voice your concerns or by representing your concerns if you feel unable to do so"

Thank you

his page asks you to look at each area in turn, for example accommodation and use the colour coded guidelines

Please tick the appropriate box below

This area remains a serious problem for me, despite any help I am given (unmet need)

'his area is not a serious problem for me because of help I am given (met need)

This area is not a serious problem for me at all (no need)

I do not want to answer this question

ccommodation - *what kind of place do you live in?*	☐	☐	☐	☐	[canu0101]
ood – *do you get enough to eat?*	☐	☐	☐	☐	[canu0102]
ooking after the home- *are you able to look after your home?*	☐	☐	☐	☐	[canu0103]
lf care – *do you have any problems keeping clean and tidy?*	☐	☐	☐	☐	[canu0104]
aytime activities – *how do you spend your day?*	☐	☐	☐	☐	[canu0105]
ysical health – *how well do you feel physically?*	☐	☐	☐	☐	[canu0106]
ychotic symptoms – *do you ever hear voices or have problems with your thoughts?*	☐	☐	☐	☐	[canu0107]
formation about condition and treatment – *have you been given clear information about your medication?*	☐	☐	☐	☐	[canu0108]
ychological distress – *have you recently felt very sad or low?*	☐	☐	☐	☐	[canu0109]
afety to self – *do ever have thoughts of harming yourself?*	☐	☐	☐	☐	[canu0110]
afety to others – *do you think you could be a danger to other people's safety?*	☐	☐	☐	☐	[canu0111]
lcohol – *does drinking cause you any problems?*	☐	☐	☐	☐	[canu0112]
rugs – *do you take any drugs that aren't prescribed?*	☐	☐	☐	☐	[canu0113]
ompany - *are you happy with your social life?*	☐	☐	☐	☐	[canu0114]
timate relationships – *do you have a partner?*	☐	☐	☐	☐	[canu0115]
xual expression – *how is your sex life?*	☐	☐	☐	☐	[canu0116]
ildcare – *do you have any children under the age of 18?*	☐	☐	☐	☐	[canu0117]
sic education – *do you have difficulty in reading, writing or understanding English?*	☐	☐	☐	☐	[canu0118]
ephone – *do you know how to use a telephone?*	☐	☐	☐	☐	[canu0119]
nsport – *how do you find using the bus, train, or tube?*	☐	☐	☐	☐	[canu0120]
ney – *how do you find budgeting your money?*	☐	☐	☐	☐	[canu0121]
efits – *are you getting all the benefits you are entitled to?*	☐	☐	☐	☐	[canu0122]

This page asks you how satisfied you are with several aspects of your life

Please answer each question by circling one number for each question below.
If there is a question you do not want to answer, leave that question blank.

1. How satisfied are you with your life as a whole today? [man0101]

1	2	3	4	5	6	7
Couldn't be worse	Displeased	Mostly Dissatisfied	Mixed	Mostly Satisfied	Pleased	Couldn't Be better

2. How satisfied are you with your job as your main occupation? (or sheltered employment or training/education) [man0102a]

1	2	3	4	5	6	7
Couldn't be worse	Displeased	Mostly Dissatisfied	Mixed	Mostly Satisfied	Pleased	Couldn't Be better

OR... if unemployed or retired...
How satisfied are you with being unemployed / retired? [man0102b]

1	2	3	4	5	6	7
Couldn't be worse	Displeased	Mostly Dissatisfied	Mixed	Mostly Satisfied	Pleased	Couldn't Be better

3. How satisfied are you with your financial situation? [man0103]

1	2	3	4	5	6	7
Couldn't be worse	Displeased	Mostly Dissatisfied	Mixed	Mostly Satisfied	Pleased	Couldn't Be better

4. How satisfied are you with the number and quality of your friendships? [man0104]

1	2	3	4	5	6	7
Couldn't be worse	Displeased	Mostly Dissatisfied	Mixed	Mostly Satisfied	Pleased	Couldn't Be better

5. How satisfied are you with your leisure activities? [man0105]

1	2	3	4	5	6	7
Couldn't be worse	Displeased	Mostly Dissatisfied	Mixed	Mostly Satisfied	Pleased	Couldn't Be better

6. How satisfied are you with your accommodation? [man0106]

1	2	3	4	5	6	7
Couldn't be worse	Displeased	Mostly Dissatisfied	Mixed	Mostly Satisfied	Pleased	Couldn't Be better

7. How satisfied are you with your personal safety? [man0107]

1	2	3	4	5	6	7
Couldn't be worse	Displeased	Mostly Dissatisfied	Mixed	Mostly Satisfied	Pleased	Couldn't Be better

8. How satisfied are you with the people that you live with? [man0108a]

1	2	3	4	5	6	7
Couldn't be worse	Displeased	Mostly Dissatisfied	Mixed	Mostly Satisfied	Pleased	Couldn't be better

OR…if you live alone…
How satisfied are you with living alone? [man0108b]

1	2	3	4	5	6	7
Couldn't be worse	Displeased	Mostly Dissatisfied	Mixed	Mostly Satisfied	Pleased	Couldn't be better

9. How satisfied are you with your sex life? [man0109]

1	2	3	4	5	6	7
Couldn't be worse	Displeased	Mostly Dissatisfied	Mixed	Mostly Satisfied	Pleased	Couldn't be better

10. How satisfied are you with your relationship with your family? [man0110]

1	2	3	4	5	6	7
Couldn't be worse	Displeased	Mostly Dissatisfied	Mixed	Mostly Satisfied	Pleased	Couldn't be better

11. How satisfied are you with your physical health? [man0111]

1	2	3	4	5	6	7
Couldn't be worse	Displeased	Mostly Dissatisfied	Mixed	Mostly Satisfied	Pleased	Couldn't be better

12. How satisfied are you with your mental health? [man0112]

1	2	3	4	5	6	7
Couldn't be worse	Displeased	Mostly Dissatisfied	Mixed	Mostly Satisfied	Pleased	Couldn't be better

This page asks you about your relationship with your staff member

Please answer each question, by marking on the ruler below.
If there is a question you do not want to answer leave that question blank.

1. Is the treatment / help you are currently receiving right for you?

0 1 2 3 4 5 6 7 8 9 10

not at all |-------|-------|-------|-------|-------|-------|-------|-------|-------|-------| **entirely**

[hasu0101]

2. Do you feel understood by your staff member?

0 1 2 3 4 5 6 7 8 9 10

not at all |-------|-------|-------|-------|-------|-------|-------|-------|-------|-------| **entirely**

[hasu0102]

3. Do you feel criticised by your staff member?

0 1 2 3 4 5 6 7 8 9 10

entirely |-------|-------|-------|-------|-------|-------|-------|-------|-------|-------| **not at all**

hasu0103]

4. Is your staff member committed to and actively involved in your treatment?

0 1 2 3 4 5 6 7 8 9 10

not at all |-------|-------|-------|-------|-------|-------|-------|-------|-------|-------| **entirely**

[hasu0104]

5. Do you trust in your staff member and in their professional competence?

0 1 2 3 4 5 6 7 8 9 10

not at all |-------|-------|-------|-------|-------|-------|-------|-------|-------|-------| **entirely**

[hasu0105]

6. How do you feel immediately after a session with your staff member?

[hasu0106]

Worse Unchanged Better

focus

Feedback of Outcome to Users and Staff

STAFF

Your name:

Please fill in the enclosed questionnaire about…

and either leave it in the FOCUS pigeon hole at your base or post it in the enclosed internal envelope addressed to:

FOCUS Study
Box P029
Institute of Psychiatry
FREEPOST LON15630
London
SE5 8BR

Direct Telephone Line: 020 7848 5095
E-mail: focus@iop.kcl.ac.uk

Thank you

This page asks you to assess the service user in each area, using the rating:

2= this is a serious problem for the service user <u>despite</u> help being given (unmet need)
1= this is not a serious problem for the service user <u>because of</u> help given (met need)
0= this is not a problem for the service user at all (no need)
9= I do not know whether this is a problem for the service user (not known)

1. Accommodation *What kind of place does the service user live in?*		[cans0101]
2. Food *Does the service user get enough to eat?*		[cans0102]
3. Looking after the home *Is the service user able to look after their home?*		[cans0103]
4. Self care *Does the service user have problems keeping clean and tidy?*		[cans0104]
5. Davtime activities *How does the service user spend their day?*		[cans0105]
6. Phvsical health *How well does the service user feel physically?*		[cans0106]
7. Psvchotic svmptoms *Does the service user ever hear voices or have problems with their thoughts?*		[cans0107]
8. Information on condition and treatment *Has the service user been given clear information about their medication?*		[cans0108]
9. Psvchological distress *Has the service user recently felt very sad or low?*		[cans0109]
10. Safetv to self *Does the service user ever have thoughts of harming themself?*		[cans0110]
11. Safetv to others *Does the service user think they could be a danger to other people's safety?*		[cans0111]
12. Alcohol *Does drinking cause the service user any problems?*		[cans0112]
13. Drugs *Does the service user take any drugs that aren't prescribed?*		[cans0113]
14. Companv *Is the service user happy with their social life?*		[cans0114]
15. Intimate relationships *Does the service user have a partner?*		[cans0115]
16. Sexual expression *How is the service user's sex life?*		[cans0116]
17. Childcare *Does the service user have any children under 18?*		[cans0117]
18. Basic education *Does the service user have any difficulty in reading, writing or understanding English?*		[cans0118]
19. Telephone *Does the service user know how to use a telephone?*		[cans0119]
20. Transport *How does the service user find using the bus, tube or train?*		[cans0120]
21. Monev *How does the service user find budgeting their money?*		[cans0121]
22. Benefits *Is the service user getting all the money they are entitled to?*		[cans0122]

THRESHOLD ASSESSMENT GRID (TAG)

THIS PAGE ASKS YOU TO RATE THE SERVICE USER'S SEVERITY OF MENTAL HEALTH PROBLEMS

For each domain (numbered 1 to 7), tick ONE statement that best applies to the person being assessed. There should be a total of 7 ticks on the completed grid (one for each domain). 'Very Severe' is only available for domains where life-saving emergency action by specialist mental health teams may be required.

		NONE	MILD	MODERATE	SEVERE	VERY SEVERE
SAFETY	**Domain 1** Intentional self harm [tagf01]	No concerns about risk of deliberate self-harm or suicide attempt ○	Minor concerns about risk of deliberate self-harm or suicide attempt ○	Definite indicators of risk of deliberate self-harm or suicide attempt ○	High risk to physical safety as a result of deliberate self-harm or suicide attempt ○	Immediate risk to physical safety as a result of deliberate self-harm or suicide attempt ○
	Domain 2 Unintentional self harm [tagf02]	No concerns about unintentional risk to physical safety ○	Minor concerns about unintentional risk to physical safety ○	Definite indicators of unintentional risk to physical safety ○	High risk to physical safety as a result of self-neglect, unsafe behaviour or inability to maintain a safe environment ○	
RISK	**Domain 3** Risk from others [tagf03]	No concerns about risk of abuse or exploitation from other individuals or society ○	Minor concerns about risk of abuse or exploitation from other individuals or society ○	Definite risk of abuse or exploitation from other individuals or society ○	Positive evidence of abuse or exploitation from other individuals or society ○	
	Domain 4 Risk to others [tagf04]	No concerns about risk to physical safety or property of others ○	Antisocial behaviour ○	Risk to property and/or minor risk to physical safety of others ○	High risk to physical safety of others as a result of dangerous behaviour ○	Immediate risk to physical safety of others as a result of dangerous behaviour ○
NEEDS AND DISABILITIES	**Domain 5** Survival [tagf05]	No concerns about basic amenities, resources or living skills ○	Minor concerns about basic amenities, resources or living skills ○	Marked lack of basic amenities, resources or living skills ○	Serious lack of basic amenities, resources or living skills ○	Life-threatening lack of basic amenities, resources or living skills ○
	Domain 6 Psychological [tagf06]l	No disabling or distressing problems with thinking, feeling or behaviour ○	Minor disabling or distressing problems with thinking, feeling or behaviour ○	Disabling or distressing problems with thinking, feeling or behaviour ○	Very disabling or distressing problems with thinking, feeling or behaviour ○	
	Domain 7 Social [tagf07]	No disabling problems with activities or in relationships with other people ○	Minor disabling problems with activities or in relationships with other people ○	Disabling problems with activities or in relationships with other people ○	Very disabling problems with activities or in relationships with other people ○	

This page asks you about your relationship with the service user

Please answer each question, by marking a point on the ruler below.
If there is a question you do not want to answer leave that question blank.

1. Do you get along with the service user?

0 1 2 3 4 5 6 7 8 9 10
not at all|------|------|------|------|------|------|------|------|------|------| **extremely well**

[hass0101]

2. Do you understand the service user and his/her views?

0 1 2 3 4 5 6 7 8 9 10
not at all|------|------|------|------|------|------|------|------|------|------| **extremely well**

[hass0102]

3. Do you look forward to meeting the service user?

0 1 2 3 4 5 6 7 8 9 10
not at all|------|------|------|------|------|------|------|------|------|------| **entirely**

[hass0103]

4. Do you feel you are actively involved in the treatment of the service user?

0 1 2 3 4 5 6 7 8 9 10
not at all|------|------|------|------|------|------|------|------|------|------| **completely**

[hass0104]

5. Do you feel you can help and effectively treat the service user?

0 1 2 3 4 5 6 7 8 9 10
not at all |------|------|------|------|------|------|------|------|------|------| **entirely**

[hass0105]

Supplementary-P

[btime]

In the box opposite please record - in minutes - how long it took the service user to complete the questionnaire

[bmanso01]

1 **Do you have anyone who you would call a "close friend"?**

0=No, 1= Yes, 9= Don't know

[bmanso02]

2 **In the last week have you seen a friend (visited a friend, been visited by a friend, or met a friend outside both your home and work)?**

0=No, 1= Yes, 9= Don't know

[bmanso03]

3 **In the past year have you been accused of a crime?**

0=No, 1= Yes, 9= Don't know

[bmanso04]

4 **In the past year have you been a victim of physical violence?**

0=No, 1= Yes, 9= Don't know

CHORD	SUPERFLUOUS
ACHE	SIMILE
DEPOT	BANAL
AISLE	QUADRUPED
BOUQUET	CELLIST
PSALM	FAÇADE
CAPON	ZEALOT
DENY	DRACHM
NAUSEA	AEON
DEBT	PLACEBO
COURTEOUS	ABSTEMIOUS
RAREFY	DÉTENTE
EQUIVOCAL	IDYLL
NAÏVE	PUERPERAL
CATACOMB	AVER
GAOLED	GAUCHE
THYME	TOPIARY
HEIR	LEVIATHAN
RADIX	BEATIFY
ASSIGNATE	PRELATE
HIATUS	SIDEREAL
SUBTLE	DEMESNE
PROCREATE	SYNCOPE
GIST	LABILE
GOUGE	CAMPANILE

[nartlang]

Baseline - National Adult Reading Test – NART (Second Edition)

Is English your first language? Yes=1 No=0

A '✓' is given if the service gives the correct answer a 'X' if the answer is incorrect

CHORD	-	SUPERFLUOUS	-
ACHE	-	SIMILE	-
DEPOT	-	BANAL	-
AISLE	-	QUADRUPED	-
BOUQUET	-	CELLIST	-
PSALM	-	FAÇADE	-
CAPON	-	ZEALOT	-zelot
DENY	-	DRACHM	-
NAUSEA	-	AEON	-
DEBT	-	PLACEBO	-
COURTEOUS	-	ABSTEMIOUS	-steee
RAREFY	-	DÉTENTE	-
EQUIVOCAL	-	IDYLL	-id (freud)
NAÏVE	-	PUERPERAL	-pwer
CATACOMB	-	AVER	-A verr
GAOLED	-	GAUCHE	-
THYME	-	TOPIARY	-toe…
HEIR	-	LEVIATHAN	-
RADIX	-raydix	BEATIFY	-be atify
ASSIGNATE	-	PRELATE	-pre-lit
HIATUS	-	SIDEREAL	-psy deereal
SUBTLE	-	DEMESNE	-demain
PROCREATE	-pro	SYNCOPE	-syncopay
GIST	-	LABILE	-
GOUGE	-	CAMPANILE	-

[nart]

Enter total number of errors in box

1. Have you been concerned about your physical health?

Have you had any physical illnesses or seen a medical doctor?

---

2. Have you felt worried or anxious?

Do unpleasant thoughts constantly go round and round in your mind?

Did you heart beat fast? (or sweat, tremble, choke?)

Has it interfered with your ability to perform your usual activities / work?

---

3. Have you felt unhappy or depressed?

How much of the time?

Are you able to switch your attention to other pleasant topics when you want to?

Have your interests in work, hobbies and social or recreational activities changed?

Has it interfered with your ability to perform your usual activities / work?

---

4. Have you been thinking about past problems?

Do you tend to blame yourself for things that happened?

Have you done anything you are still ashamed of?

---

5. How have you been getting along with people (family, friends, co-workers?)

Have you been irritable or grumpy lately?

Have you been involved in any fights?

---

6. Do you ever feel uncomfortable as if people were watching you?

Is anyone trying to harm or interfere with you in any way?

Are you concerned about anybody's intentions towards you?

Have you felt that people are out to get you?

---

7. Have things or events had special meaning for you? Have you see any references to yourself on TV or in the newspapers?

Do you have a special relationship with God?

How do you explain the things that have been happening?

Have you felt that were under the control of another person or force?

---

8. Is there a special purpose or mission to your life?

Do you have any special powers or abilities?

Have you thought that you might be somebody rich or famous?

---

9. Have you heard any sounds or people talking to you or about you when there has been nobody around?

Have you seen any visions or smelled any smells others don't seem to notice?

Have these experiences interfered with your ability to perform your usual activities / work?

---

10. May I ask you one or two standard questions we ask everybody?

How old are you? What is the date?

What is this place called? / Where are you?

Baseline - Brief Psychiatric Rating Scale

INSTRUCTIONS:
There are 18 items to be rated. Items 11-18 should be rated on the basis of observations made during the interview. For these items, 1 = Not observed. The remaining items should be rated on the basis of reported (i.e. subjective) information pertaining to the past week. For these items, 1 = Not reported.
Please insert appropriate number in corresponding box.

bprs01]

1. **SOMATIC CONCERN:** Degree of concern over present bodily health. Rate the degree to which physical health is perceived as a problem by the patient, whether complaints have a realistic basis nor not. *Do Not* rate mere reporting of somatic symptoms. Rate only concern for (or worrying about) physical problems (real of imagined).
0 = Cannot be assessed adequately because of severe formal thought disorder, uncooperativeness, or marked evasiveness/guardedness, or; Not assessed.
1 = Not reported.
2 = Very Mild: occasionally is somewhat concerned about body, symptoms, or physical illness.
3 = Mild: occasionally is moderately concerned, or often is somewhat concerned.
4 = Moderate: occasionally is very concerned, or often is moderately concerned.
5 = Moderately Severe: often is very concerned.
6 = Severe: is very concerned most of the time.
7 = Very Severe: is very concerned nearly all of the time.

prs02]

2. **ANXIETY:** Worry, fear, or over-concern for present or future. Rate solely on the basis of verbal report of patient's own subjective experiences. Do not infer anxiety from physical signs or from physical signs or from neurotic defence mechanisms. **Do not rate if restricted to somatic concern.**
0 = Cannot be assessed adequately because of severe formal thought disorder, uncooperativeness, or marked evasiveness/guardedness, or; Not assessed.
1 = Not reported.
2 = Very Mild: occasionally is somewhat anxious.
3 = Mild: occasionally fees moderately anxious, or often feels somewhat anxious.
4 = Moderate: occasionally is very anxious, or often feels moderately anxious.
5 = Moderately Severe: often is very anxious.
6 = Severe: feels very anxious most of the time.
7 = Very Severe: feels very anxious nearly all of the time.

rs03]

3. **DEPRESSIVE MOOD:** Subjective report of feeling depressed, blue, `down in the dumps', etc. Rate only degree of reported depression. *Do Not* rate on the basis of inferences concerning depression based upon general retardation and somatic complaints.
0 = Cannot be assessed adequately because of severe formal thought disorder, uncooperativeness, or marked evasiveness/guardedness, or; Not assessed.
1 = Not reported.
2 = Very Mild: occasionally feels somewhat depressed.
3 = Mild: occasionally feels moderately depressed, or often feels somewhat depressed.
4 = Moderate: occasionally feels very depressed.
5 = Moderately Severe: often feels very depressed.
6 = Severe: feels very depressed most of the time.
7 = Very Severe: feels very depressed nearly all of the time.

bprs04]

4. **GUILT FEELINGS:** Over-concern or remorse for past behaviour. Rate on the basis of the patient's subjective experiences of guilt as evidenced by verbal report. Do not infer guilt feelings from depression, anxiety or neurotic defences.
0 = Cannot be assessed adequately because of severe formal thought disorder, uncooperativeness, or marked evasiveness/guardedness, or; Not assessed.
1 = Not reported.
2 = Very Mild: occasionally feels somewhat guilty.
3 = Mild: occasionally feels moderately guilty, or often feels somewhat guilty.

4 = Moderate: occasionally feels very guilty, or often feels moderately guilty.
5 = Moderately Severe: often feels very guilty.
6 = Severe: feels very guilty most of the time, or encapsulated delusion of guilt.
7 = Very Severe: agonising constant feelings of guilt, or pervasive delusion(s) of guilt.

---

bprs05]

5. **HOSTILITY:** Animosity, contempt, belligerence, disdain for other people outside the interview situation. Rate solely on the basis of the verbal report of feelings and actions of the patient towards others; *Do Not* infer hostility from neurotic defences, anxiety or somatic complaints.
0 = Cannot be assessed adequately because of severe formal thought disorder, uncooperativeness, or marked evasiveness/guardedness, or; Not assessed.
1 = Not reported.
2 = Very Mild: occasionally feels somewhat angry.
3 = Mild: often feels somewhat angry, or occasionally feels moderately angry.
4 = Moderate: occasionally feels very angry, or often feels moderately angry.
5 = Moderately Severe: often feels very angry.
6 = Severe: has acted on his anger by being verbally or physically abusive on one or two occasions.
7 = Very Severe: has acted on his anger on several occasions.

ɔrs06]

6. **SUSPICIOUSNESS:** Belief (delusional or otherwise) that others have now, or have had in the past, malicious or discriminatory intent towards the patient. On the basis of verbal report, rate only those suspicions which are currently held whether they concern past or present circumstances.
0 = Cannot be assessed adequately because of severe formal thought disorder, uncooperativeness, or marked evasiveness/guardedness, or; Not assessed.
1 = Not reported.
2 = Very Mild: rare instances of distrustfulness which may or may not be warranted by the situation.
3 = Mild: occasional instances of suspiciousness that are definitely not warranted by the situation.
4 = Moderate: more frequent suspiciousness, or transient ideas of reference.
5 = Moderately Severe: pervasive suspiciousness, or frequent ideas of reference.
6 = Severe: definite, delusion(s) of reference or persecution that is (are) not wholly pervasive (e.g. an encapsulated delusion).
7 = Very Severe: as above, but more widespread, frequent, or intense.

ɔ07]

7. **UNUSUAL THOUGHT CONTENT:** Severity of delusions of any type – consider conviction, and effect on actions. Assume full conviction if patient has acted on his or her beliefs.
0 = Cannot be assessed adequately because of severe formal thought disorder, uncooperativeness, or marked evasiveness/guardedness, or; Not assessed.
1 = Not reported.
2 = Very Mild: delusion(s) suspected or likely.
3 = Mild: at times, patient questions his or her belief(s) (partial delusion).
4 = Moderate: full delusional conviction, but delusion(s) has little or no influence on behaviour.
5 = Moderately Severe: full delusional conviction, but delusion(s) has only occasional impact on behaviour.
6 = Severe: delusion(s) has significant effect, e.g. neglects responsibilities because of preoccupations with belief that he-she is God.
7 = Very Severe: delusion(s) has major impact, e.g. stops eating because believes food is poisoned.

---

ors08]

8. **GRANDIOSITY:** Inflated self-esteem (self-confidence), or inflated appraisal of one's talents, powers, abilities, accomplishments, knowledge, importance, or identity. *Do Not* score mere grandiose *quality* of claims (e.g. `I'm the worst sinner in the world', `The entire country is trying to kill me') unless the guilt/persecution is related to some special, exaggerated attributes: e.g. if patient denies talents, powers etc, even if he or she states that *others* indicate that he/she has these attributes, this item should not be scored.

0 = Cannot be assessed adequately because of severe formal thought disorder, uncooperativeness, or marked evasiveness/guardedness, or; Not assessed.
1 = Not reported.
2 = Very Mild: e.g. is more confident that most people, but of only possible clinical significance.
3 = Mild: definitely inflated self-esteem or exaggerates talents somewhat out of proportion to the circumstances.
4 = Moderate: e.g. inflated self-esteem clearly out of proportion to the circumstances, or suspect grandiose delusion(s).
5 = Moderately Severe: e.g. single (definite) encapsulated grandiose delusion, or multiple (definite) fragmentary grandiose delusions.
6 = Severe: e.g. a single (definite grandiose delusion/delusional system, or multiple (definite) grandiose delusions that the patient seems preoccupied with.
7 = Very Severe: e.g. as above, but nearly all conversation is directed towards the patient's grandiose delusion(s).

---

rs09]

9. **HALLUCINATORY BEHAVIOUR:** Perceptions (in any sensory modality) in the absence of an identifiable external stimulus. Rate only those experiences that have occurred during the last week. *Do Not* rate `voices in my head' or `visions in my mind' unless the patient can differentiate between the experiences and his or her thoughts.

0 = Cannot be assessed adequately because of severe formal thought disorder, uncooperativeness, or marked evasiveness/guardedness, or; Not assessed.
1 = Not reported.
2 = Very Mild: suspected hallucinations only.
3 = Mild: definite hallucinations, but insignificant, infrequent, or transient (e.g. occasional formless visual hallucinations, a voice calling the patient's name).
4 = Moderate: as above, but more frequent or extensive (e.g. frequently sees the devil's face, two voices carry on lengthy conversations).
5 = Moderately Severe: hallucinations are experienced nearly every day, or are a source of extreme distress.
6 = Severe: as above, and has had a moderate impact on the patient's behaviour (e.g. concentration difficulties leading to impaired work functioning).
7 = Very Severe: as above, and has had a severe impact (e.g. attempts suicide in response to command hallucinations).

:10]

10. **DISORIENTATION:** Confusion or lack of proper association for person, place or time.

0 = Cannot be assessed adequately because of severe formal thought disorder, uncooperativeness, or marked evasiveness/guardedness, or; Not assessed.
1 = Not observed.
2 = Very Mild: e.g. seems somewhat confused.
3 = Mild: e.g. indicates 1982 when, in fact, it is 1983.
4 = Moderate: e.g. indicates 1978.
5 = Moderately Severe: e.g. is unsure where he/she is.
6 = Severe: e.g. has no idea where he/she is.
7 = Very Severe: e.g. does not know who he/she is.

THE FOLLOWING ITEMS SHOULD BE COMPLETED AFTER THE INTERVIEW

)bprs11]

11. **EMOTIONAL WITHDRAWAL:** Deficiency in rating to the interviewer and to the interview situation. Overt manifestations of this deficiency include poor/absence of eye control, failure to orient oneself physically toward the interviewer, and a general lack of involvement or engagement in the interview. Distinguish from *Blunted Affect*, in which deficits in facial expression, body gesture, and voice pattern are scored.
    0 = Cannot be assessed (e.g. scored from audiotape).
    1 = Not observed.
    2 = Very Mild: e.g. occasionally exhibits poor eye contact.
    3 = Mild: e.g. as above, but more frequent.
    4 = Moderate: e.g. exhibits little eye contact, but still seems engaged in the interview and is appropriately responsive to all questions.
    5 = Moderately Severe: e.g. stares at floor or orients self away from interviewer, but still seems moderately engaged.
    6 = Severe: e.g. as above, but more persistent or pervasive.
    7 = Very Severe: e.g. appears `spacey' or `out of it' (total absence of emotional relatedness), and is disproportionately uninvolved or unengaged in the interview. (*Do not score if explained by disorientation).*

)prs12]

12. **CONCEPTUAL DISORGANISATION:** Degree of speech incomprehensibility. Include any type of formal thought disorder (e.g. loose associations, incoherence, flight of ideas, neologisms). *Do Not* include mere circumstantiality or pressured speech, even if marked. *Do Not* rate on the basis of the patient's subjective impressions (e.g. `my thoughts are racing. I can't hold a thought', `my thinking gets all mixed up'). Rate *only* on the basis of observations made during the interview.
    1 = Not observed.
    2 = Very Mild: e.g. somewhat vague, but of doubtful clinical significance.
    3 = Mild: e.g. frequently vague, but the interview is able to progress smoothly.
    4 = Moderate: e.g. occasional irrelevant statements, infrequent use of neologisms, or moderate loosening of associations.
    5 = Moderately Severe: as above, but more frequent.
    6 = Severe: formal thought disorder is present for most of the interview, and the interview is severely strained.
    7 = Very Severe: very little coherent information can be obtained.

)rs13]

13. **TENSION:** Rate motor restlessness (agitation) observed during the interview. *Do Not* rate on the basis of subjective experiences reported by the patient. Disregard suspected pathogenesis (e.g. tardive dyskinesia).
    0 = Cannot be assessed (e.g. scored from audiotape).
    1 = Not observed.
    2 = Very Mild: e.g. occasionally fidgets.
    3 = Mild: e.g. frequently fidgets.
    4 = Moderate: e.g. constantly fidgets, or frequently fidgets, wrings hands and pulls clothing.
    5 = Moderately Severe: e.g. constantly fidgets, wrings hands and pulls at clothing.
    6 = Severe: e.g. cannot remain seated (i.e. must pace).
    7 = Very Severe: e.g. paces in a frantic manner.

rs14]

14. **MANNERISMS AND POSTURING:** Unusual and unnatural motor behaviour. Rate only abnormality of movements; do not rate simple heightened motor activity here. Consider frequency, duration, and degree of bizarreness. Disregard suspected pathogenesis.
    0 = Cannot be assessed (e.g. scored from audiotape).
    1 = Not observed.

2 = Very Mild: odd behaviour but of doubtful clinical significance, e.g. occasional unprompted smiling, infrequent lip movements.
3 = Mild: strange behaviour but not obviously bizarre, e.g. infrequent head tilting (side to side) in a rhythmic fashion, intermittent abnormal finger movements.
4 = Moderate: e.g. assumes yoga position for a brief period of time, infrequent tongue protrusions, rocking.
5 = Moderately Severe: e.g. unusual movements in several body areas.
6 = Severe: as above, but more frequent, intense, or pervasive.
7 = Very Severe: e.g. bizarre posturing throughout most of the interview, Continuous abnormal movements in several body areas.

bprs15]

15. **MOTOR RETARDATION:** Reduction in energy level evidenced in slowed movements. Rate on the basis of observed behaviour of the patient only. *Do Not* rate on the basis of the patient's subjective impression of his or her own energy level.
0 = Cannot be assessed (e.g. scored from audiotape)
1 = Not observed.
2 = Very Mild and of doubtful clinical significance.
3 = Mild: e.g. conversation is somewhat retarded, movements somewhat slowed.
4 = Moderate: e.g. conversation is somewhat retarded, movements somewhat slowed.
5 = Moderately Severe: e.g. conversation is strained, moves very slowly.
6 = Severe: e.g. conversation is difficult to maintain, hardly moves at all.
7 = Very Severe: e.g. conversation is almost impossible, does not move at all throughout the interview.

s16]

16. **UNCOOPERATIVENESS:** Evidence of resistance, unfriendliness, resentment, and lack of readiness to cooperate with the interviewer. Rate only on the basis of the patient's attitude and responses to the interviewer and the interview situation. *Do Not* rate on the basis of reported resentment or uncooperativeness outside the interview situation.
1 = Not observed.
2 = Very Mild: e.g. does not seem motivated.
3 = Mild: e.g. seems evasive in certain areas.
4 = Moderate: e.g. monosyllabic, fails to elaborate spontaneously.
5 = Moderately Severe: e.g. expresses resentment and is unfriendly throughout the interview.
6 = Severe: e.g. refuses to answer a number of questions.
7 = Very Severe: e.g. refuses to answer most questions.

s17]

17. **BLUNTED AFFECT:** Diminished affective responsively, as characterised by deficits in facial expression, body gesture, and voice pattern. Distinguish from *Emotional Withdrawal*, in which the focus is on interpersonal impairment rather than affect. Consider degree and consistency of impairment.
0 = Cannot be assessed (e.g. scored from audiotape)
1 = Not observed.
2 = Very Mild: e.g. occasionally seems indifferent to material that is usually accompanied by some show of emotion.
3 = Mild: e.g. somewhat diminished facial expression, *or* somewhat monotonous voice *or* somewhat restricted gestures.
4 = Moderate: e.g. as above, but more intense, prolonged, or frequent.
5 = Moderately Severe: e.g. flattening of affect, including at leave *two* or the three features: severe lack of facial expression, monotonous voice, or restricted body gestures.
6 = Severe: e.g. profound flattening of affect.
7 = Very Severe: e.g. total monotonous voice, *and* total lack of expressive gestures throughout the evaluation.

18. **EXCITEMENT:** Heightened emotional tone, including irritability and expansiveness (hypomanic affect). *Do Not* infer affect from statements of grandiose delusions.

Baseline - Brief Psychiatric Rating Scale

1 = Not observed.
2 = Very Mild and of doubtful clinical significance.
3 = Mild: e.g. irritable or expansive at times.
4 = Moderate: e.g. frequently irritable or expansive.
5 = Moderately Severe: e.g. constantly irritable or expansive; or, at times, enraged or euphoric.
6 = Severe: e.g. enraged or euphoric throughout most of the interview.
7 = Very Severe: e.g. as above, but to such a degree that the interview must be terminated prematurely.

Baseline Supplementary-S

[btimes]

In the box opposite please record - in minutes - how long it took the staff member to complete the questionnaire

What are the positive aspects of your relationship with your patient?

[bpos]

What are the negative aspects of your relationship with the patient?

[bneg]

Baseline - Health of the Nation Outcome Scale

Summary of rating instructions

1 Rate each scale in order from 1 to 12, place the number in the box above which applies best to the service user.

2 Do not include information rated in an earlier item except for item 10 which is an overall rating.

3 Rate the MOST SEVERE problem that has occurred during the previous 2 weeks.

4 All scales follow the same format:

0 = no problem
1 = minor problem requiring no action
2 = mild problem but definitely present
3 = moderately severe problem
4 = severe or very severe problem
Rate 9 if not known

HoNOS Glossary

[bhonos01]

1. Overactive, aggressive, disruptive or agitated behaviour.

- *Include such behaviour due to any cause (e.g. drugs, alcohol, dementia, psychosis, depression, etc.).*

Do not include bizarre behaviour rate at Scale 6

0	No problem of this kind during the period rated.
1	Irritability, quarrels, restlessness, etc. not requiring action.
2	Includes aggressive gestures, pushing or pestering others; threats or verbal aggression; lesser damage to property (e.g. broken cup, window); marked overactivity or agitation.
3	Physically aggressive to others or animals (short of rating 4); threatening manner; more serious overactivity or destruction of property.
4	At least one serious physical attack on others or on animals; destructive of property (e.g. fire-setting); serious intimidation or obscene behaviour.

[bhonos02]

2. Non-accidental self-injury.

- *Do not include accidental self-injury (e.g. due to dementia or severe learning disability); the cognitive problem is rated at Scale 4 and the injury at Scale 5.*
- *Do not include illness or injury as a direct consequence of drug /alcohol use, rated at Scale 3 (e.g. cirrhosis of the liver or injury resulting from drink driving are rated at Scale 5).*

0	No problem of this kind during the period rated.
1	Fleeting thoughts about ending it all but little risk during period rated; no self-harm.
2	Mild risk during period rated; includes non-hazardous self-harm (e.g. wrist scratching).
3	Moderate to serious risk of deliberate self-harm during period rated; includes preparatory acts (e.g. collecting tablets).
4	serious suicidal attempt and /or serious deliberate self-injury during period rated.

[bhonos03]

3. Problem-drinking or drug taking.

- *Do not include aggressive / destructive behaviour due to alcohol or drug use, rated at Scale 1.*
- *Do not include physical illness or disability due to alcohol or drug use, rated at Scale 5.*

0	No problem of this kind during the period rated.
1	Some over indulgence but within social norm.
2	Loss of control of drinking or drug taking, but not seriously addicted.
3	Marked cravings or dependence on alcohol or drugs with frequent loss of control, risk taking under the influence.
4	Incapacitated by alcohol / drug problem.

[bhonos04]

4. Cognitive problems.

- *Include problems of memory, orientation and understanding associated with any disorder: learning disability, dementia, schizophrenia etc.*
- *Do not include temporary problems (e.g. hangovers) resulting from drugs / alcohol use, rated at Scale 3.*

0 No problem of this kind during the period rated.
1 Minor problems with memory or understanding (e.g. forgets names occasionally).
2 Mild but definite problems (e.g. has lost the way in a familiar place or failed to recognise a familiar person); sometimes mixed up about simple decisions.
3 Marked disorientation in time, place or person; bewildered by everyday events; speech is sometimes incoherent; mental slowing
4 Severe disorientation (e.g. unable to recognise relatives); at risk of accidents; speech incomprehensible; clouding or stupor.

[bhonos05]

5. Physical illness or disability problems.

- *Include illness or disability from any cause that limits or prevents movement, or impairs sight or hearing, or otherwise interferes with personal functioning.*
- *Include side-effects from medication; effects of drug/ alcohol use; physical disabilities resulting from accidents or self-harm associated with cognitive problems, drink-driving, etc.*
- *Do not include mental or behavioural problems rated at Scale 4.*

0 No physical health problem during the period rated.
1 Minor health problems during the period (e.g. cold, non-serious fall).
2 Physical health problem imposes mild restriction on mobility and activity.
3 Moderate degree of restriction on activity due to physical health problem.
4 Severe or complete incapacity due to physical health problem.

[bhonos06]

6. Problems associated with hallucinations and delusions.

- *Include hallucinations and delusions irrespective of diagnosis.*
- *Include odd and bizarre behaviour associated with hallucinations or delusions.*
- *Do not include aggressive, destructive or overactive behaviours attributed to hallucinations or delusions rated at Scale 1.*

0 No evidence of hallucinations or delusions during period rated.
1 Somewhat odd or eccentric beliefs not in keeping with cultural norms.
2 Delusions or hallucinations (e.g. voices, visions) are present, but there is little distress to patient or manifestation in bizarre behaviour.
3 Marked preoccupation with delusions or hallucinations, causing much distress and/ or manifested in obviously bizarre behaviour, i.e. moderately severe clinical problem.
4 Mental state and behaviour is seriously and adversely affected by delusions or hallucinations, with severe impact on patient.

bhonos07]

7. Problems with depressed mood.

- *Do not include overactivity or agitation, rated at Scale 1.*
- *Do not include suicidal ideation of attempts, rated at Scale 2*
- *Do not include delusions or hallucinations, rated at Scale 6*

0 No problem associated with depressed mood during period rated.
1 Gloomy; or minor changes in mood
2 Mild but definite depression and distress (e.g. feelings of guilt; low self-esteem).
3 Depression with inappropriate self-blame; preoccupied with feelings of guilt.
4 Severe or very severe depression, with guilt or self accusation.

[bhonos08]

8. Other mental and behavioural problems.

- *Rate only the most severe clinical problem not considered at items 6 and 7 as follows.*
- *Specify the rate of the problem by entering the appropriate letter: A phobic; B anxiety; C obsessive-compulsive; D mental strain/ tension; E dissociative; F somatoform; G eating; H sleep; I sexual; J other, specify.*

0 No evidence of any of these problems during period rated.
1 Minor health problems only.
2 A problem is clinically present at a mild level (e.g. patient has a degree of control).
3 Occasional severe attacks of distress, with loss of control (e.g. has to avoid anxiety provoking situations altogether, call in a neighbour to help etc.), i.e. moderately severe level of problem.
4 Severe problem dominates most activities.

[bhonos09]

9. Problems with relationships.

- *Rate the patients most severe problem associated with active or passive withdrawal from social relationships, and or non-supportive, destructive or self-damaging relationships.*

0 No significant problem during the period rated.
1 Minor non-clinical problem.
2 Definite problem in making or sustaining supportive relationships; patient complains and /or problems are evident to others.
3 Persisting major problems due to active or passive withdrawal from social relationships and /or to relationships that provide little or no comfort or support.
4 Severe and distressing social isolation due to inability to communicate socially and /or withdrawal from social relationships.

[bhonos10]

10. Problems with activities of daily living.

- *Rate the overall level of functioning in activities of daily living (ADL's), (e.g. problems with basic activities of self-care such as eating, washing, dressing, toilet; also complex skills such as budgeting, organising where to live. Do not include lack of opportunities for exercising intact abilities and skills, rated at Scales 11-12.*
- *Include any lack of motivation for using self-help opportunities, since this contributes to a lower overall level of functioning.*
- *Do not include lack of opportunities for exercising intact abilities and skills, rate at Scales 11-12.*

0 No problem during period rated; good ability to function in all areas.
1 Minor problems only (e.g. untidy, disorganised).
2 Self-care adequate, but major lack of performance of one or more complex skills (see above).
3 Major problem in one or more area of self-care (see above) as well as major inability to perform several complex skills.
4 Severe disability or incapacity in all or nearly all areas of self-care and complex skills.

[bhonos11]

11. Problems with living conditions.

- *Rate the overall severity of problems with the quality of living conditions and daily domestic routine.*
- *Are the basic necessities met (heat, light, hygiene)? If so, is there help to cope with disabilities and a choice of opportunities to use skills and develop new ones?*
- *Do not rate the level of functional disability itself, rated at Scale 10.*

NB: Rate patients usual accommodation. If in an acute ward, rate the home accommodation. If information not available, rate 9.

0 Accommodation and living conditions are acceptable; helpful in keeping any disability rated at Scale 10 to the lowest level possible, and supportive of self-help.

1 Accommodation is reasonably acceptable although there are minor or transient problems (e.g. not ideal location, not preferred option, doesn't like the food etc.).

2 Significant problem with one or more aspects of the accommodation and /or regime (e.g. restricted choice; staff or household have little understanding of how to limit disability or how to help use or develop new or intact skills).

3 Distressing multiple problems with accommodation (e.g. some basic necessities absent); housing environment has minimal or no facilities to improve patient's independence.

4 Accommodation is unacceptable (e.g. lack of basic necessities, patient at risk of eviction, or roofless, or living conditions are otherwise intolerable, making patient's problems worse.

[bhonos12

12. Problems with occupation and activities.

- *Rate the overall level of problems with quality of day-time environment. Is there help to cope with disabilities, and opportunities for maintaining or improving occupational and recreational skills and activities? Consider factors such as stigma, lack of qualified staff, access to supportive facilities (e.g. staffing and equipment of day centres, workshops, social clubs, etc.).*
- *Do not rate the level of functional disability itself, rated at Scale 10.*

NB: Rate patient's usual situation. If in acute ward, rate activities during period before admission. If information not available rate 9.

0 Patient's day-time environment is acceptable; helpful in keeping any disability rated at Scale 10 to the lowest level possible, and supportive of self-help.

1 Minor or temporary problems (e.g. late giro cheques); reasonable facilities available but not always at desired times etc.

2 Limited choice of activities; lack of reasonable tolerance (e.g. unfairly refused entry to public library or baths etc.); handicapped by lack of permanent address; insufficient carer or professional support; helpful day setting available but for very limited hours.

3 Marked deficiency in skilled services available to help minimise level of existing disability; no opportunities to use intact skills or add new ones; unskilled care difficult to access

4 Lack of any opportunity for day-time activities makes patients problem worse.

Sociodemographic Form

PIN:	Date:	Researcher: - MG – LM - AP – MS

Section 1 - Contact Details

First Name	
Last Name	
Date of Birth	

Current address OR **Contact address**	

Telephone Number	

Sociodemographic Form

PIN:	

[pin]

Sex
Insert appropriate number in box

Sex [ ]
1=Male
2=Female

[sdsex]

Ethnicity
Insert appropriate number in box below

1=White
2=Black Caribbean
3=Black African
4=Black other
5=Indian
6=Pakistani
7=Bangladeshi
8=Chinese
9=Other Asian
10=Other

[ ]
[sdeth]

Community Mental Health Team
Insert appropriate number in box

Community Mental Health Team allocated to
1=North West Croydon
2=North North Croydon
3=North East Croydon
4=West Central Croydon
5=Mid Central Croydon
6=Central East Croydon
7=South West Croydon
8=South East Croydon

[ ]
[sdcmht]

Mental Health Service Contact

/ /	Date of first contact for **this** period of care

[sdpoc]

/ /	Date of **first ever contact** with mental health services (in months/years)

[sdfc]

[ ]	**Please tick box** if date of first ever contact with mental health services (in months/years) is **NOT KNOWN**

[sdfckn]

Sociodemographic Form

Clinical Diagnosis
Insert appropriate number in box

Primary Clinical Diagnosis

[sdcd]

Secondary Clinical Diagnosis

[sdcd]

Education
Insert appropriate number in box below

Age on first leaving full time education
1=<17
2=17-18
3=19 – 22
4=> 23
9 = NOT KNOWN

[sdaed]

Highest Educational Level
1=No formal Qualifications
2=GCSE / GCE or equivalent
3=A levels or equivalent
4=Higher diploma or Degree
5=Post-graduate degree
9=Not Known

[sdhel]

Casenote form

PIN:	Date:	Researcher: - MG – LM - AP – MS -

Direct Support – please tick relevant boxes (current information only)

Support for Service User

- Psychological ☐ [bcdup]
- Social ☐ [bcdus]
- Emotional ☐ [bcdue]

Monitoring

- Needs ☐ [bcdmn]
- Risk ☐ [bcdmr]
- Medication compliance ☐ [bcdmmc]

Psychological Support

- Supportive counseling ☐ [bcdpsc]
- Cognitive/ behavioural ☐ [bcdpcb]
- Psychodynamic ☐ [bcdpp]
- Eclectic ☐ [bcdpe]

Pharmacological

Exact Prescription

[bmedctn]

Depot ☐ 1=Yes 0=No [bcdpd]

Casenote form

***Assessments**

Psychological ☐ [bcdap]

Neurological ☐ [bcdan]

Specialist Unit ☐ [bcdasu]

Other ☐ [bcdao] Please specify [bcdaosp]

*** Information from the previous 18 months only**

Indirect Support – please tick relevant boxes (current information only)

Liaison with other agencies

GP ☐ [bcidgp]

Social Services ☐ [bcidss]

Other ☐ [bcidoth] Please specify [bcidoths]

Carer Support

Focus on Service User's needs

Psychological ☐ [bcidcup]

Social ☐ [bcidcus]

Emotional ☐ [bcidcue]

Focus on Carer's needs

Psychological ☐ [bcidccp]

Social ☐ [bcidccs]

Emotional ☐ [bcidcce]

Is the service user involved in any other research project?
0 = No 1 = Yes

☐ [bcnr]

Please specify name or type of research project

[bcnrne]

SERVICE USER Impact of involvement questionnaire

Please answer all the questions in order.

0 = No
1 = Yes

Did **filling in the FOCUS questionnaires** make you think about the care you are getting? [ ] [fupfq01]

yes, in what way?

[fupfq01i]

Did **filling in the FOCUS questionnaires** make you think about your relationship with your staff member? [ ] [fupfq02]

yes, in what way?

[fupfq02i]

Did you receive the feedback? [ ] [fupfq03]

Did you read the feedback? [ ] [fupfq04]

Did you understand the feedback? [ ] [fupfq05]

Did **receiving the feedback** make you think about the care you are receiving? [ ] [fupfq06]

es, in what way?

[fupfq06i]

Did **receiving the feedback** make you think about your relationship with your staff member? [ ] [fupfq07]

es, in what way?

[fupfq07i]

d receiving the feedback lead you to discuss the content of your care with your staff member? [ ] [fupfq08]

id receiving the feedback lead you to change your behaviour with your staff member [ ] [fupfq09]

e.g. by discussing your relationship with them?

Do you have any comment on the questionnaires or feedback? *e.g.* the content of the questionnaires, the at of the feedback, the frequency of the questionnaire (monthly) or the feedback (3-monthly)

[fupfq10i]

STAFF Impact of Involvement questionnaire

Please answer all the questions in order.

0 = No
1 = Yes

Did **filling in the FOCUS questionnaires** make you think about the care the service user gets? ☐ [fupfq01]

yes, in what way?

[fupfq01i]

Did **filling in the FOCUS questionnaires** make you think about your relationship with the service user? ☐ [fupfq02]

yes, in what way?

[fupfq02i]

Did you receive the feedback? ☐ [fupfq03]

Did you read the feedback? ☐ [fupfq04]

Did you understand the feedback? ☐ [fupfq05]

Did **receiving the feedback** make you think about the care the service user is receiving? ☐ [fupfq06]

es, in what way?

[fupfq06i]

Did **receiving the feedback** make you think about your relationship with the service user? ☐ [fupfq07]

es, in what way?

[fupfq07i]

id receiving the feedback lead you to discuss the content of their care with the service user? ☐ [fupfq08]

id receiving the feedback lead you to change your behaviour with the service user ☐ [fupfq09]

e.g. by discussing your relationship with them?

Do you have any comment on the questionnaires or feedback? *e.g.* the content of the questionnaires, the at of the feedback, the frequency of the questionnaire (monthly) or the feedback (3-monthly)

[fupfq10i]

Adverse Event Form

PIN:	Date:	Researcher: - MG – LM - AP – MS

Adverse event categories:	
1. Suicide	
2. Attempted Suicide	
3. Homicide	
4. Violent Crime	
5. Other	

Advevent

Details: (eg: Source of information, event pre- or post- allocation, etc)

Admission Form

PIN:

Is the service user in hospital at time of entry into the study?

[admentryl]

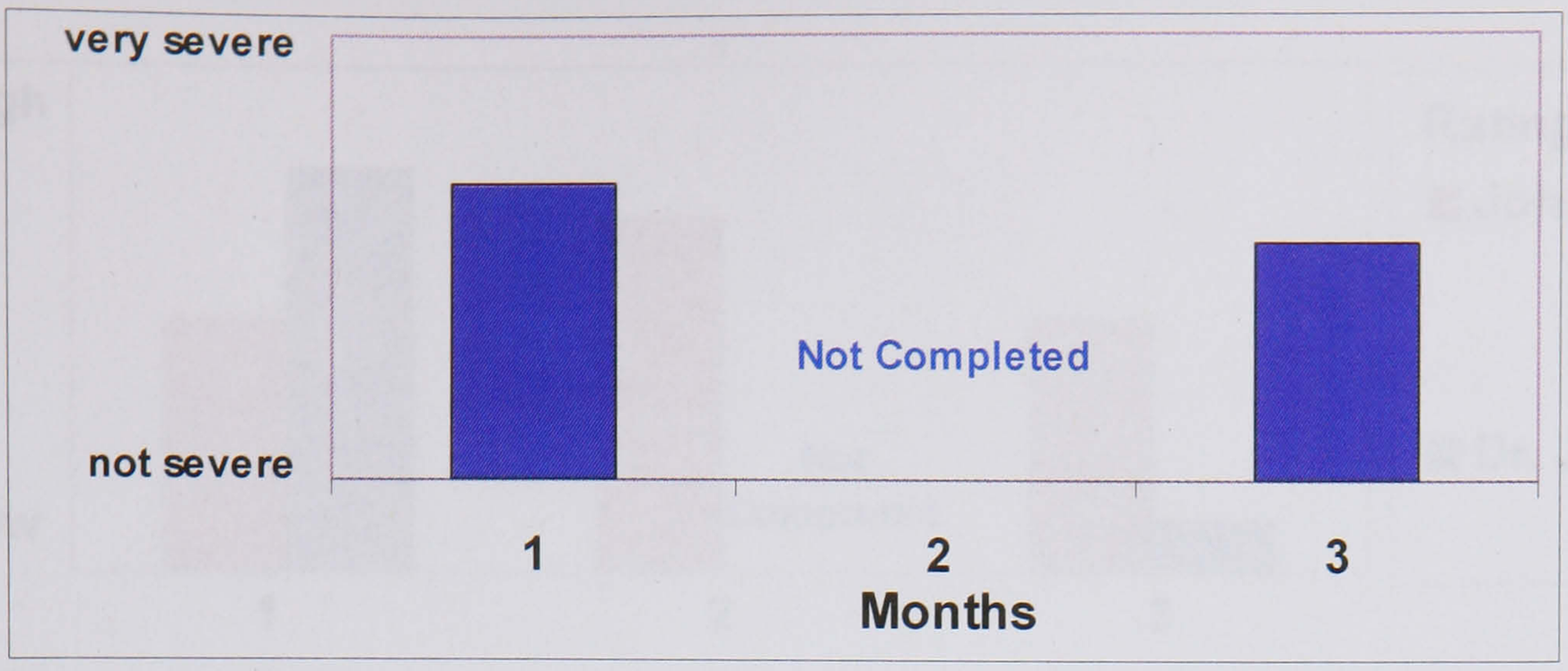
Yes = 1 **No=0**

Please enter number of days for each admission from entry to study (date of second baseline interview either with staff member or service user)

No. of admissions	Name of Hospital	No. of days as a psychiatric in-patient	No. of days as a non-psychiatric in-patient
1			
[admit101]	[admit102]	[admit103]	[admit104]
2			
[admit201]	[admit202]	[admit203]	[admit204]
3			
[admit301]	[admit302]	[admit303]	[admit304]
4			
[admit401]	[admit402]	[admit403]	[admit404]
5			
[admit501]	[admit502]	[admit503]	[admit504]
6			
[admit601]	[admit602]	[admit603]	[admit604]
7			
[admit701]	[admit702]	[admit703]	[admit704]
8			
[admit801]	[admit802]	[admit803]	[admit804]
9			
[admit901]	[admit902]	[admit903]	[admit904]
10			
[admit1001]	[admit1002]	[admit1003]	[admit1004]

NOTE: An in-patient day includes an overnight stay

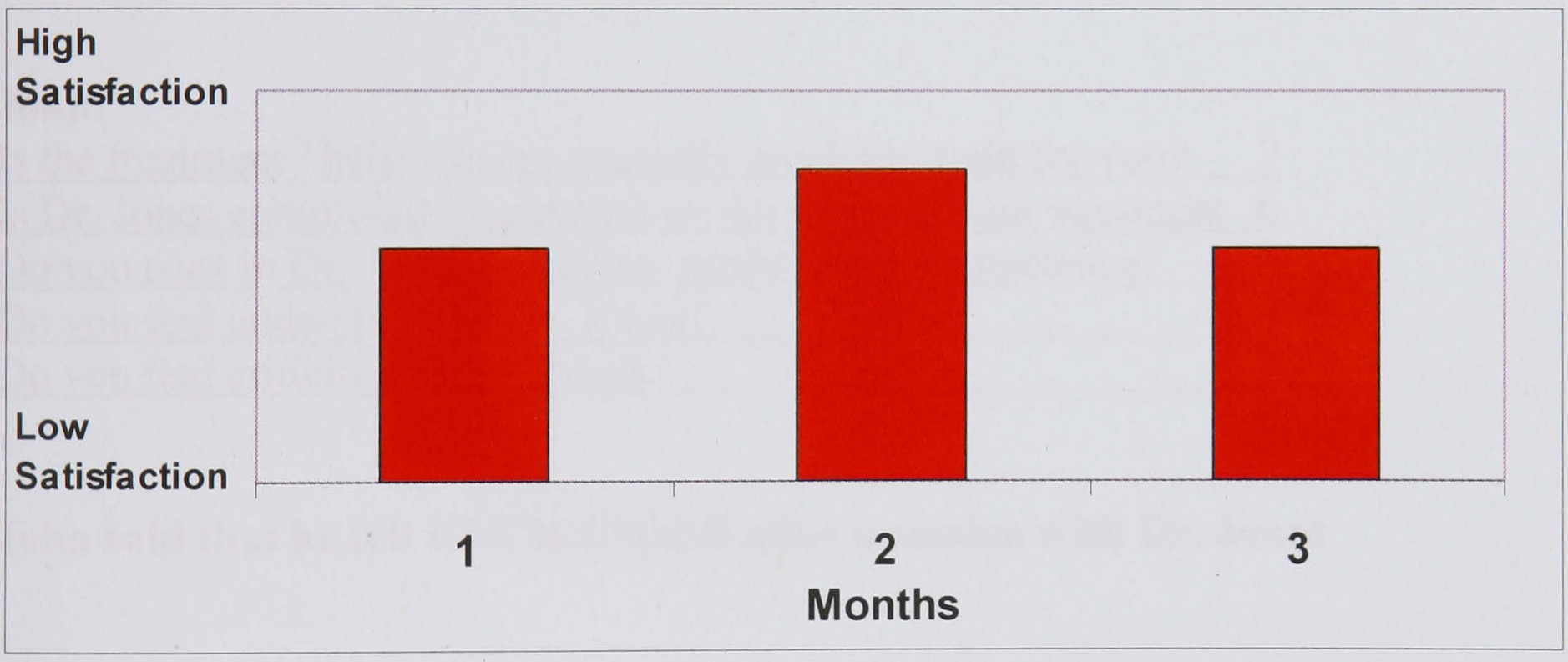
Severity of mental health problems rated by Dr. Jones over the past 3 months



The last ratings by Dr. Jones were:

Intentional self harm	Moderate
Unintentional self harm	Moderate
Risk from others	Moderate
Risk to others	Moderate
Survival	Severe
Psychological	Severe
Social	Severe

Quality of life rated by John



The last ratings by John were:

Life today	Mixed
Employment / unemployment	Mostly Dissatisfied
Financial situation	Mostly Dissatisfied
Number and quality of friendships	Mostly Dissatisfied
Leisure activities	Mostly Dissatisfied
Accommodation	Mostly Dissatisfied
Personal safety	Mostly Dissatisfied
People that you live with	Mixed
Sex life	Mostly Satisfied
Relationship with family	Mixed
Physical health	Mostly satisfied
Mental health	Mixed

Last ratings
(Ratings go from 0 = not at all to 10 = extremely well)

Dr. Jones:

Do you get along with John?	8
Do you understand John and his views?	8
Do you look forward to meeting with John?	6
Do you feel you can help and effectively treat John?	3
Do you feel you are actively involved in the treatment of John?	2

John:

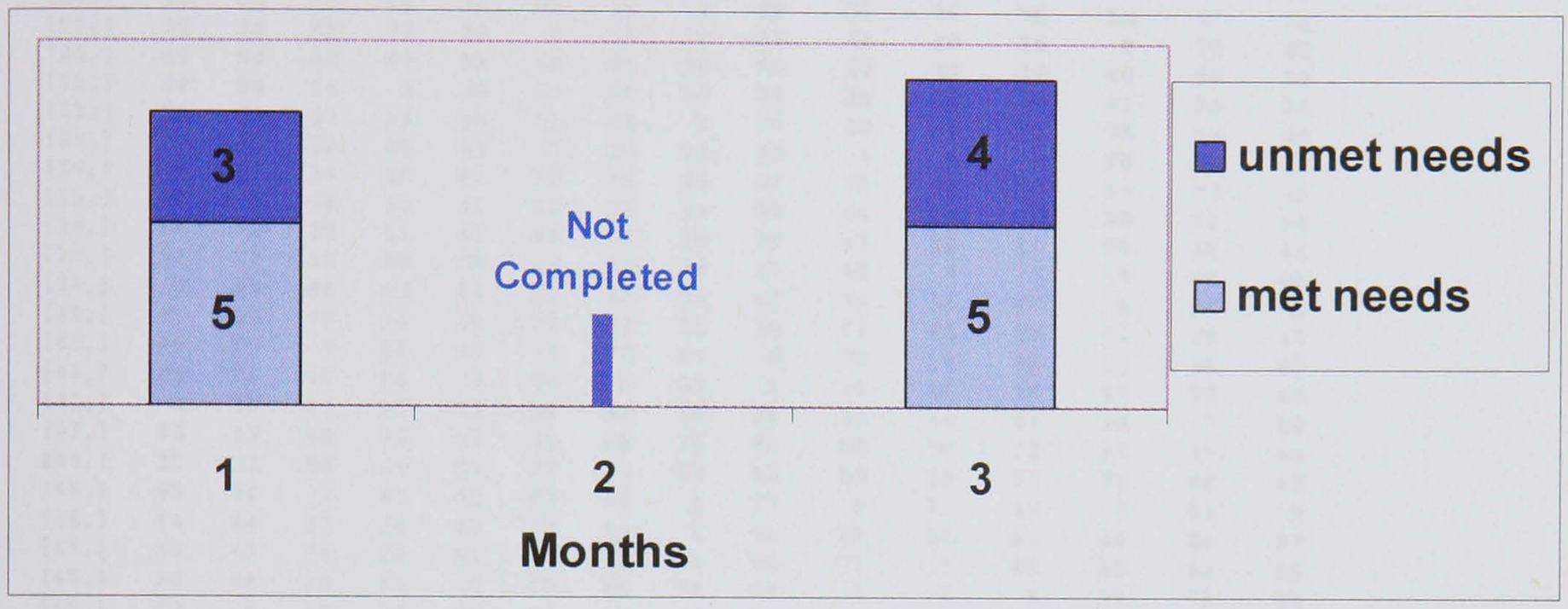
Is the treatment / help you are currently receiving right for you?	7
Is Dr. Jones committed to and actively involved in your treatment	6
Do you trust in Dr. Jones and in her professional competence?	6
Do you feel understood by Dr. Jones?	5
Do you feel criticised by Dr. Jones?	5

John said that he felt UNCHANGED after a session with Dr. Jones

Met & Unmet Need

Assessment by John

Assessment by Dr. Jones



Areas of Agreement in the last assessment

Unmet needs
Accommodation
Benefits
Food
Daytime activities

Met needs
Looking after the home
Self care

Areas of Disagreement

Rated as unmet need by John:
Money (rated as met by Dr. Jones)
Safety to self (rated as met by Dr. Jones)

Rated as met need by John:
Sexual Expression (rated as not known by Dr. Jones)
Intimate Relationships (rated as no need by Dr. Jones)

Rated as No Need by John
Childcare (rated as unknown by Dr. Jones)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
[1,]	63	54	36	33	50	36	84	32	59	51	23	31	59	54	75
[2,]	32	6	63	76	66	61	47	12	5	34	25	8	47	48	22
[3,]	24	73	8	38	54	16	41	27	42	2	37	56	17	76	48
[4,]	52	13	6	82	68	6	58	7	27	27	51	48	86	84	57
[5,]	79	67	3	78	77	45	57	60	51	15	71	7	8	17	17
[6,]	42	83	41	18	86	87	26	21	74	68	85	70	2	35	13
[7,]	47	60	4	19	58	1	7	8	40	75	41	46	50	42	68
[8,]	30	44	13	29	69	80	82	61	30	67	86	29	26	16	54
[9,]	56	19	65	39	21	14	50	48	71	38	48	19	82	41	8
[10,]	61	35	26	48	56	49	19	1	21	78	34	74	49	83	84
[11,]	67	23	54	2	18	77	79	87	22	39	33	42	90	21	61
[12,]	76	12	47	71	9	85	4	11	19	46	29	37	39	19	90
[13,]	72	55	67	73	33	66	31	15	20	82	28	33	38	6	2
[14,]	64	69	55	20	46	27	38	86	62	90	38	16	70	72	87
[15,]	28	4	81	45	73	5	40	49	2	16	76	5	69	70	50
[16,]	74	49	59	23	90	42	3	58	76	63	84	34	41	59	65
[17,]	23	66	83	79	79	37	64	38	56	8	60	57	53	38	16
[18,]	33	7	1	7	71	41	43	31	11	61	27	13	73	43	21
[19,]	81	17	35	10	5	38	5	85	9	14	55	52	85	61	23
[20,]	19	29	16	63	1	18	23	81	79	44	32	26	19	63	14
[21,]	2	14	51	90	74	8	83	82	8	36	2	64	22	1	7
[22,]	78	21	84	27	75	46	65	36	85	60	15	38	1	75	19
[23,]	51	77	72	8	6	7	69	64	86	21	49	10	62	56	35
[24,]	88	36	88	41	26	48	54	67	18	41	31	27	52	87	27
[25,]	9	75	38	9	83	30	75	20	57	17	39	55	35	58	25
[26,]	73	34	75	1	17	29	33	19	90	31	74	51	33	69	66
[27,]	5	20	52	57	59	82	48	88	69	83	54	14	76	23	63
[28,]	53	51	76	66	72	55	36	2	88	56	26	76	12	44	12
[29,]	35	56	45	70	24	9	55	89	16	24	80	35	9	25	42
[30,]	45	62	89	89	29	65	81	30	52	40	46	22	45	52	72
[31,]	68	58	14	3	39	57	90	62	78	30	16	80	13	20	52
[32,]	50	89	53	13	30	51	76	5	65	10	43	72	75	65	69
[33,]	71	28	20	46	67	3	24	28	39	4	5	4	28	31	44
[34,]	65	41	34	55	82	52	18	68	67	76	63	30	14	73	33
[35,]	27	1	58	52	45	54	34	24	50	64	56	32	20	12	64
[36,]	66	5	90	51	42	63	25	35	24	89	22	11	56	30	46
[37,]	31	72	21	25	61	12	12	76	37	42	9	78	15	60	86
[38,]	3	45	80	60	64	11	22	54	83	74	44	67	6	32	1
[39,]	41	25	70	26	10	72	29	51	34	13	45	24	51	78	40
[40,]	40	80	9	85	57	71	70	44	6	70	19	66	11	81	85
[41,]	82	70	49	86	3	64	8	23	1	11	18	54	27	79	60
[42,]	1	38	17	65	53	22	52	56	25	22	24	87	23	7	20
[43,]	12	33	48	72	62	44	68	72	82	88	79	12	87	37	49
[44,]	37	53	56	64	52	17	30	80	12	59	59	77	71	68	18
[45,]	87	74	2	40	49	83	16	9	17	6	72	15	7	64	9
[46,]	11	64	37	34	43	32	51	6	41	23	20	61	44	24	67
[47,]	36	47	24	81	51	60	13	74	14	73	1	85	10	82	55
[48,]	22	48	18	61	4	47	46	73	64	5	6	9	37	71	43
[49,]	13	9	29	15	87	21	2	41	63	49	69	83	63	9	82
[50,]	26	71	64	36	28	2	72	55	10	45	67	18	25	3	29
[51,]	38	11	12	59	88	68	9	18	60	43	89	68	61	18	36
[52,]	43	88	62	69	44	59	78	50	72	18	8	2	29	11	74
[53,]	60	31	32	37	19	31	66	77	73	57	7	65	57	74	26
[54,]	85	82	86	30	80	33	28	70	46	29	75	84	42	50	62
[55,]	69	61	43	87	13	79	20	46	3	65	50	63	40	45	41
[56,]	58	86	7	43	11	50	39	42	54	81	90	88	24	49	71
[57,]	59	18	44	56	16	88	71	34	35	35	11	36	18	88	73
[58,]	15	3	73	49	12	34	80	78	15	32	21	25	66	2	70
[59,]	77	8	78	50	55	67	86	4	89	50	52	75	5	89	80
[60,]	80	16	79	84	81	81	35	17	23	55	57	41	46	62	10
[61,]	4	10	31	16	32	43	59	22	75	80	83	44	79	5	32
[62,]	84	26	25	28	31	20	44	79	29	48	73	53	60	77	5
[63,]	21	24	22	44	25	35	60	10	80	3	3	1	21	86	39
[64,]	18	15	87	14	36	70	87	84	58	25	53	23	32	14	31
[65,]	6	76	71	68	38	86	6	52	13	28	65	50	65	53	4
[66,]	57	81	42	77	70	24	89	43	43	19	17	17	77	85	30
[67,]	10	32	5	80	37	40	32	66	48	77	40	60	74	13	83
[68,]	62	39	30	53	41	53	11	39	55	12	14	39	48	8	78
[69,]	55	57	57	22	22	19	10	90	36	79	77	59	64	67	6
[70,]	86	87	39	83	20	76	15	65	28	20	68	40	16	47	89
[71,]	29	59	19	42	34	90	63	29	77	52	42	62	84	40	24
[72,]	90	68	27	17	78	15	73	83	38	54	78	79	36	39	76
[73,]	48	27	69	21	8	84	14	37	68	85	88	3	67	46	34
[74,]	14	79	10	5	7	89	67	69	4	9	30	86	43	4	28
[75,]	83	46	33	4	63	75	62	14	49	71	87	73	31	22	77
[76,]	75	85	15	62	85	4	17	57	31	86	12	45	54	33	15
[77,]	49	2	50	47	23	58	21	33	7	69	62	69	78	90	11
[78,]	17	84	61	58	84	56	27	47	26	1	47	43	80	27	3
[79,]	70	37	28	75	47	28	1	16	45	66	66	20	72	28	37
[80,]	7	42	11	31	2	10	37	26	70	7	58	81	55	15	81
[81,]	25	63	85	24	60	78	45	45	84	62	81	47	4	34	38
[82,]	89	43	82	6	35	73	49	13	61	37	13	49	88	55	79
[83,]	39	40	66	12	15	13	77	53	53	53	82	28	34	51	45
[84,]	8	30	60	88	76	62	61	40	33	87	35	89	3	29	47
[85,]	46	50	23	54	89	25	74	25	44	26	61	21	81	80	58
[86,]	54	78	46	67	14	69	53	3	66	47	36	90	89	36	51
[87,]	44	22	40	74	40	23	88	63	47	84	4	82	58	26	88
[88,]	20	65	74	32	48	74	56	75	32	58	10	71	83	10	59
[89,]	34	90	77	35	27	39	85	59	81	33	70	6	30	66	56
[90,]	16	52	68	11	65	26	42	71	87	72	64	58	68	57	53

Appendix 2

This Appendix contains publications arising from the FOCUS Study:

Slade M, Leese M, Ruggeri M, Kuipers E, Tansella M, Thornicroft G (2004) *Does meeting needs improve quality of life?*, Psychotherapy and Psychosomatics, **73**, 183-189.

Slade M (2002) *What outcomes to measure in routine mental health services, and how to assess them – a systematic review*, Australian and New Zealand Journal of Psychiatry, **36**, 743-753.

Slade M (2002) *The use of patient-level outcomes to inform treatment*, Epidemiologia e Psichiatria Sociale, **11**, 20-27.

Slade M (2002) *Routine outcome assessment in mental health services*, Psychological Medicine, **32**, 1339-1344.

Psychotherapy and Psychosomatics

Regular Article

Psychother Psychosom 2004;73:183–189
DOI: 10.1159/000076456

Does Meeting Needs Improve Quality of Life?

Mike Slade[a] Morven Leese[a] Mirella Ruggeri[b] Elizabeth Kuipers[c]
Michele Tansella[b] Graham Thornicroft[a]

[a]Health Services Research Department, Institute of Psychiatry, London, UK; [b]Department of Medicine and Public Health, University of Verona, Verona, Italy; [c]Department of Psychology, Institute of Psychiatry, London, UK

Key Words
Mental health · Quality of life · Outcome assessment (health care)

Abstract

Background: This study investigated the relationship between patient-rated unmet needs and subjective quality of life using routine outcome data. ***Methods:*** 265 mental health service patients from South Verona were assessed using the Camberwell Assessment of Need, the Lancashire Quality of Life Profile, and other standardised assessments of symptoms, disability, function and service satisfaction. At 1-year follow-up, 166 patients were still in contact, of whom 121 patients (73%) were reassessed. ***Results:*** Higher baseline quality of life was associated with being male, a diagnosis of psychosis, higher disability, higher satisfaction with care, fewer staff-rated or patient-rated unmet needs, and fewer patient-rated met needs (accounting for 40% of the variance). Specifically, fewer baseline patient-rated unmet needs were cross-sectionally associated with a higher quality of life (B = –0.08, 95% CI –0.12 to –0.04). Apart from its baseline value, the only baseline predictor of follow-up QoL was patient-rated unmet need (B = –0.08, 95% CI –0.21 to –0.09), accounting for 58% of the variance in follow-up quality of life. Graphical chain modelling confirmed this association. ***Conclusions:*** The association between high numbers of unmet needs and low subjective quality of life appears increasingly robust across several studies. Future research will need to investigate whether changes in needs precede changes in quality of life. This study provides further evidence that a policy of actively assessing and addressing patient-rated unmet needs may lead to improved quality of life.

Introduction

'Need' in mental health care includes broader areas of health and social functioning than most other areas of medicine. Needs can be rated by staff or patients, and have been differentiated into unmet needs (current serious problems, whether or not help is offered or given) and met needs (no or moderate problems due to help given) [1]. There is a growing consensus across Europe that mental health care should be provided on the basis of need, with an intended goal of improving subjective quality of life [2, 3]. This goal is feasible if patients with the most needs have the lowest subjective quality of life, and if

KARGER

Fax + 41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

0033–3190/04/0733–0183$21.00/0

Accessible online at:
www.karger.com/pps

Dr. Mike Slade, MRC Clinician Scientist Fellow
Box P029, Institute of Psychiatry
De Crespigny Park, Denmark Hill
London, SE5 8AF (UK)
Tel. +44 20 7848 0795, Fax +44 20 7277 1462, E-Mail m.slade@iop.kcl.ac.uk

meeting needs (i.e. changing them from unmet needs to met needs) *causes* subjective quality of life to improve. This study investigated whether a causal relationship exists between needs and subjective quality of life.

There is preliminary but robust evidence that patient-rated unmet needs are cross-sectionally associated with lower subjective quality of life. The UK700 study assessed staff-rated but not patient-rated needs, and found that patient-rated quality of life was predicted more by unmet needs than any other clinical or social variable [4]. A study from the PRiSM Psychosis Study compared the relative contribution of staff and patient assessments of needs, and found that number of unmet needs was inversely associated with quality of life (replicating the UK700 study), and that the association was more robust for patient ratings of need [5]. Current evidence therefore indicates that patient-rated unmet need may be particularly influential on quality of life. Both these studies employed researchers to collect the outcome data – cross-sectional association has not been investigated using routinely-collected outcome data.

Bollen [6] proposed three criteria for establishing a causal relationship: association (the putative cause and effect have temporal and spatial contiguity), direction (cause precedes effect) and isolation (the effects of a cause are isolated from competing causes). Using this framework, there is currently cross-sectional evidence for association and isolation in the relationship between patient-rated unmet need and quality of life, but no longitudinal evidence whether meeting patient-rated unmet needs *causes* subjective quality of life to improve.

Hypotheses

Two hypotheses were investigated using routinely-collected outcome data:

1 The number of patient-rated unmet needs is cross-sectionally inversely associated with subjective quality of life for all patients, whether or not clinical and social variables are controlled for (cross-sectional isolation and association, respectively).
2 The number of patient-rated unmet needs at baseline predicts level of subjective quality of life at one-year follow-up (longitudinal association).

Method

Participants

The study took place in South Verona (population 75,000), a predominantly urban area in north-east Italy. The South Verona Community Mental Health Service provides comprehensive integrated services, emphasising continuity of care by employing staff (excluding nurses) who work across hospital and community facilities [7]. Data presented here are part of the larger South-Verona Outcome Project (SVOP), a naturalistic longitudinal study which assesses the outcome of care provided by the South-Verona Community Mental Health Service using standardised instruments completed within routine clinical practice. One of the major aims of the SVOP is testing, in the context of a 'real-world service', hypotheses which are related to clinically relevant issues [8, 9]. The study involves assessments of all patient in contact with the Service, with repeated follow-up assessments of regular attenders. The analysis reported here is based on the attenders assessed in one of the SVOP waves in 1996 and the sub-group who were still in contact one year later.

Procedures

Mental health staff completed four assessments at baseline and follow-up. The Global Assessment of Functioning (GAF) scale is a single-item measure of functioning from 0 (extremely severe dysfunction) to 90 (extremely good function) [10]. The Brief Psychiatric Rating Scale (BPRS) expanded version is a 24-item measure of symptomatology, covering anxiety/depression, positive symptoms, negative symptoms, mania and cognitive impairment [11]. Each item is rated from 1 (no symptom) to 7 (extremely severe symptom). The Disability Assessment Schedule (DAS) is an 8-item assessment of social role functioning, with each item rated from 0 (no dysfunction) to 5 (maximum dysfunction) [12]. The Camberwell Assessment of Need (CAN) assesses the presence of a met or unmet need in 22 health and social domains [1]. Patients also completed the CAN at baseline only. In addition, patients completed at baseline and follow-up the Verona Service Satisfaction Scale (VSSS) which assesses 54 aspects of care from 1 (terrible) to 5 (excellent) [13], and the Lancashire Quality of Life Profile (LQL) which assesses quality of life [14]. Only LQL subjective measures are reported in this study, which use a 7-point Likert scale from 1 ('My life couldn't be worse') to 7 ('My life couldn't be better') for general well-being and eight more specific domains – leisure/participation, religion, finances, living situation, legal and safety, family and social relations, health and self-concept. The mean score for all nine domains was used as the LQL score, as a meaningful indicator of subjective well-being. VSSS and LQL assessments relate to the previous year, and all other assessments to the previous month.

Statistical Analysis

Non-responder differences were tested using independent-sample t tests. Hypothesis 1 was tested using linear regression with variables entered in blocks comprising sociodemographic data (sex, age, and dichotomous variables for being employed [as opposed to unemployed, home maker, student or retired] and being married), baseline diagnosis (psychosis or not), and baseline-dependent variables (BPRS, GAF, DAS, and either staff CAN-unmet and CAN-met, patient CAN-unmet and CAN-met, or both). Hypothesis 2 was tested using linear regression analysis on baseline values (with and without baseline LQL included as an independent variable), with follow-up LQL score as the dependent variable. Percentages of variance are adjusted R^2 statistics. Hypothesis 2 was also investigated using graphical modelling [15]. A graphical chain model of the relationship between baseline and follow-up scores for CAN and LQL was constructed to illustrate the strongest relationships among all the variables taken together. Baseline variables were fixed within the model. The stepwise backward procedure was used to select a model, remov-

Table 1. Clinical and social characteristics of all patients (n = 265) and long-term patient subgroup (n = 121) at baseline

	All patients	Long-term patients
Age, mean ± SD	45.7 ± 15.5	45.8 ± 15.8
Male	95 (36%)	43 (36%)
Marital status		
Unmarried	106 (40%)	55 (46%)
Married	107 (40%)	43 (36%)
Widowed/separated/divorced	52 (20%)	23 (18%)
Living situation		
Alone	38 (14%)	18 (15%)
With family or relatives	215 (81%)	96 (79%)
Hospital/hostel	12 (5%)	7 (6%)
Employment		
Employed	97 (37%)	34 (28%)
Unemployed	36 (14%)	24 (20%)
Home-maker/retired/student	132 (50%)	63 (52%)
Diagnosis		
Schizophrenia and other functional psychosis	75 (28%)	49 (41%)
Affective psychosis	19 (7%)	9 (7%)
Depressive neurosis	94 (36%)	38 (31%)
Other neurosis	37 (14%)	11 (9%)
Personality disorder	20 (8%)	10 (8%)
Other or no psychiatric diagnosis	20 (8%)	4 (3%)
Mean assessment score ± SD		
GAF (range 0–90)	59.0 ± 15.5	56.0 ± 15.6
DAS (range 0–5)	0.63 ± 0.92	0.80 ± 1.01
BPRS (range 1–7)	1.51 ± 0.47	1.57 ± 0.53
Staff CAN-unmet (range 0–22)	0.88 ± 1.54	1.12 ± 1.60
Staff CAN-met (range 0–22)	2.37 ± 2.29	2.68 ± 2.52
Patient CAN-unmet (range 0–22)	1.22 ± 2.03	1.50 ± 2.36
Patient CAN-met (range 0–22)	1.85 ± 2.03	2.10 ± 2.12
VSSS (range 1–5)	3.95 ± 0.51	3.92 ± 0.54
LQL (range 1–7)	4.55 ± 0.85	4.49 ± 0.90

ng partial correlations not significant at p = 0.01. Graphical modeling relies on the assumptions that non-linear relationships are negligible and that the relationship between variable pairs is not modified by a third variable. 'Leave-one-out' residuals from the final fitted model were examined for evidence of non-normality and non-linearity [15]. Regression analyses were performed using SPSS for Windows 8.0.1, and estimation and fitting of the graphical model using MIM.

Results

The clinical and social characteristics of the 265 patients assessed at baseline and of the sub-group of 121 long-term users assessed at follow-up are shown in table 1.

At 1-year follow-up, assessments were attempted only for the 166 patients still in contact with the service. Twenty-three patients were too unwell to interview, 22 had follow-up staff assessment but refused to complete the self-administered instruments, giving a sample of 121 patients (73% of patients still in contact, 46% of baseline sample) with full baseline and follow-up assessments. Compared with the 121 patients with full data, the 144 patients for whom complete follow-up data were unavailable were more likely to be employed ($t = 2.7$, $p < 0.01$), have a non-psychotic diagnosis ($t = 3.7$, $p < 0.01$), and have higher GAF ($t = 3.0$, $p < 0.01$), lower DAS ($t = 2.8$, $p < 0.01$), lower BPRS ($t = 2.1$, $p = 0.04$), and higher staff CAN-unmet ($t = 2.4$, $p = 0.02$), CAN-met ($t = 2.0$, $p = 0.05$) and patient CAN-unmet ($t = 10.7$, $p < 0.01$) ratings. There was no difference in quality of life.

Table 2. Regression analysis for quality of life of all patients, using staff and patient CAN assessments of need separately and combined (n = 265)

	Staff CAN only			Patient CAN only			Combined CAN		
	B	Beta	p	B	Beta	p	B	Beta	p
Age	−0.01	−1.00	0.096	0.00	−0.07	0.200	0.00	−0.09	0.119
Sex	*−0.26*	*−0.15*	*0.005*	−0.15	−0.08	0.100	*−0.20*	*−0.11*	*0.027*
Marital status	0.16	0.09	0.083	0.16	0.09	0.084	0.15	0.09	0.089
Employment status	−0.13	−0.07	0.225	−0.09	−0.05	0.341	−0.14	−0.08	0.161
Psychosis	*0.33*	*0.18*	*0.001*	*0.28*	*0.16*	*0.004*	*0.28*	*0.16*	*0.003*
DAS	*0.23*	*0.25*	*0.004*	*0.21*	*0.23*	*0.005*	*0.27*	*0.29*	*0.001*
BPRS	−0.08	−0.05	0.608	−0.31	−0.17	0.040	−0.16	−0.09	0.311
GAF	0.01	0.13	0.146	0.01	0.13	0.118	0.01	0.11	0.196
VSSS	*0.76*	*0.45*	*<0.001*	*0.73*	*0.43*	*<0.001*	*0.65*	*0.39*	*0.000*
Staff CAN-unmet	*−0.16*	*−0.29*	*<0.001*				*−0.13*	*−0.24*	*0.001*
Staff CAN-met	−0.04	−0.12	0.068				−0.01	−0.04	0.552
Patient CAN-unmet				*−0.09*	*−0.22*	*<0.001*	*−0.08*	*−0.19*	*<0.001*
Patient CAN-met				*−0.09*	*−0.22*	*<0.001*	*−0.09*	*−0.21*	*<0.001*

Italic = $p < 0.05$.
B = Regression coefficient; Beta = coefficient when dependent and independent variable standardised to have unit standard deviation.

Table 3. Regression analysis of follow-up quality of life for long-term patients (n = 121), using baseline variables with and without baseline quality of life included

Variable	Excluding baseline LQL			Including baseline LQL		
	B	Beta	p	B	Beta	p
DAS	0.32	0.34	0.013	0.05	0.05	0.650
BPRS	0.21	0.13	0.327	0.33	0.19	0.062
GAF	*0.02*	*0.33*	*0.012*	0.01	0.16	0.144
VSSS	0.28	0.17	0.053	−0.07	−0.04	0.570
Staff CAN-unmet	−0.10	−0.18	0.112	−0.06	−0.11	0.213
Staff CAN-met	0.01	0.04	0.669	0.02	0.06	0.449
Patient CAN-unmet	*−0.15*	*−0.41*	*<0.001*	*−0.08*	*−0.23*	*0.002*
Patient CAN-met	*−0.08*	*−0.19*	*0.033*	−0.02	−0.04	0.585
LQL				*0.62*	*0.63*	*<0.001*

Italic = $p < 0.05$.
B = Regression coefficient; Beta = coefficient when dependent and independent variable standardised to have unit standard deviation.

Hypothesis 1: Cross-Sectional Isolation and Association

At baseline the non-parametric correlation for all 265 patients of LQL with patient-rated unmet needs was 0.34 ($p < 0.001$). The regression analysis on LQL for all patients is shown in table 2.

The model comprising solely sociodemographic predictors accounted for 1% of variance in LQL, with diagnosis added still accounted for 1%, and with all variables except CAN added accounted for 30% (not shown). Adding the staff CAN gave a model accounting for 34%, adding the patient CAN gave a model accounting for 37%, and adding both staff and patient CAN gave a model accounting for 40% of the variance in LQL. Higher quality of life was associated with being male, having a diagnosis of psychosis, having higher disability, having higher

satisfaction with care, having fewer unmet needs (however rated), and fewer patient-rated met needs. For the combined model, the CAN predictors were patient-rated met needs (B = –0.09, 95% CI –0.13 to –0.05), patient-rated unmet needs (B = –0.08, 95% CI –0.12 to –0.04), and staff-rated unmet need (B = –0.13, 95% CI –0.21 to –0.05). These regression coefficients imply that, for example, a reduction of one patient-rated unmet need was associated with an average increase of 0.09 in LQL score (scale 1 to 7). The same analysis undertaken on the 121 long-term patients (not shown) also found that patient-rated unmet needs were inversely associated with quality of life (B = –0.09, 95% CI –0.15 to –0.03).

Hypothesis 2: Longitudinal Association

The mean staff CAN-unmet rating at 1-year follow-up was 0.99 (a reduction of 0.12), CAN-met was 3.12 (an increase of 0.44), and of LQL was 4.63 (an increase of 0.14). The results of a regression using baseline data on follow-up LQL for the long-term patients are shown in table 3.

The model accounted for 33% of the variance in quality of life at follow-up with baseline LQL excluded, and 58% with baseline LQL included. Apart from baseline LQL, the best predictor of follow-up quality of life was baseline patient-rated unmet need (B = –0.08, 95% CI –0.21 to –0.09 excluding baseline LQL, B = –0.08, 95% CI –0.14 to –0.03 including baseline LQL).

The relationship between baseline and follow-up values for CAN and LQL was elaborated using graphical chain modelling, shown in figure 1. This technique provides a simpler, visual depiction of the structure of the relationship between variables, by setting non-significant partial correlations to [15]. The result is a graph with nodes denoting variables, edges indicating associations between variables, and the absence of an edge indicating conditional independence (although the two variables may be marginally independent if they are connected indirectly through other nodes).

When controlling for all possible relationships between variables (a feature of graphical modelling), the follow-up quality of life score was predicted by baseline patient CAN-unmet and LQL scores. In addition, the follow-up staff CAN assessments were predicted by baseline levels, and follow-up staff unmet needs were predicted by baseline staff met needs. Follow-up staff met needs were predicted by baseline staff unmet needs using all data, but this relationship became insignificant when one outlier identified by analysis of residuals) was omitted.

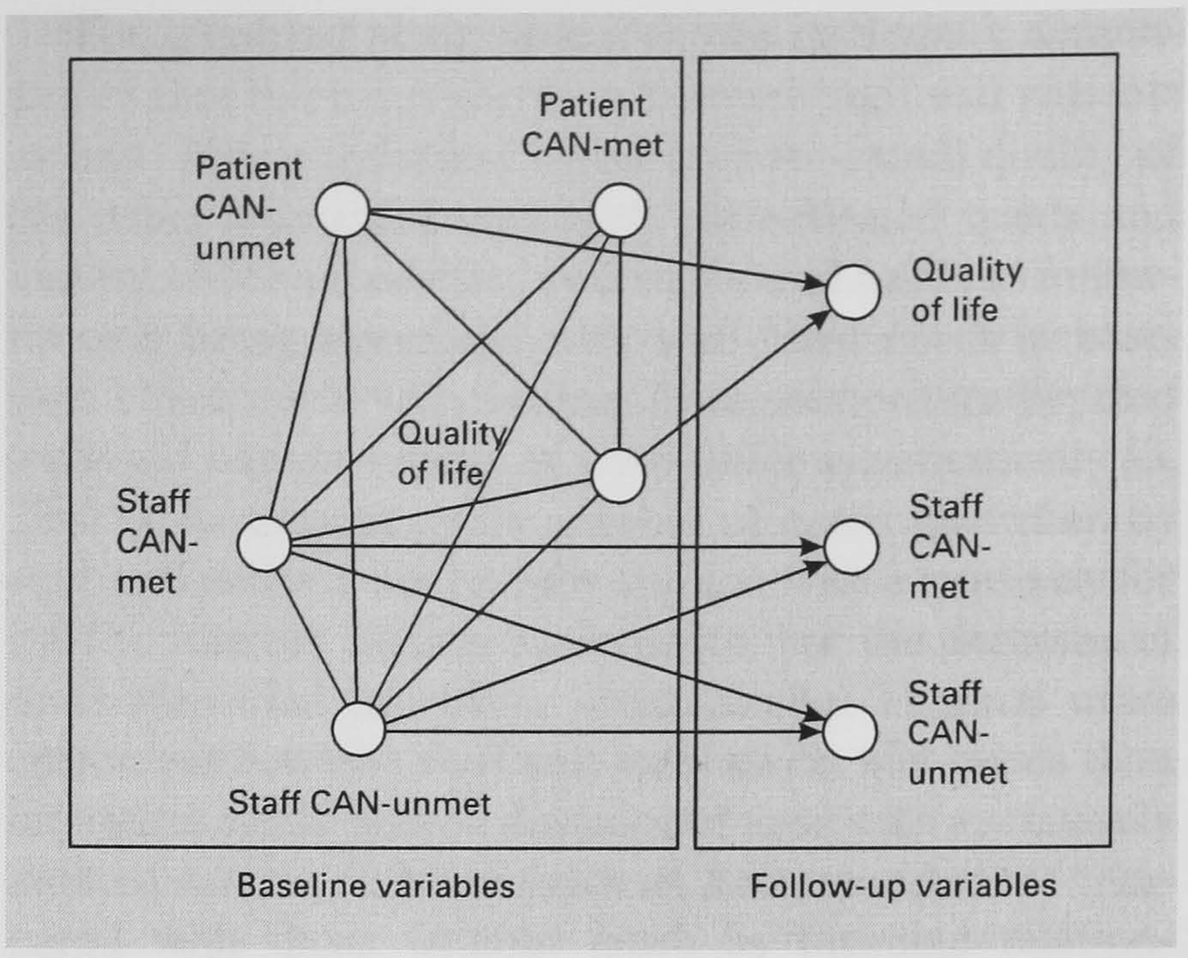


Fig. 1. Graphical chain model of baseline and follow-up needs and quality of life variables (n = 121).

Discussion

This study represents the first attempt to move beyond establishing cross-sectional associations, by investigating whether level of unmet need is temporally associated with level of subjective quality of life, and use of routine outcome data maximises its generalisability to other mental health services. Patient-rated unmet need, and to a lesser extent patient-rated met need, was cross-sectionally associated with subjective quality of life, controlling for other sociodemographic and clinical variables. The level of subjective quality of life at 1-year follow-up was also predicted by baseline level of patient-rated unmet need, whether or not baseline subjective quality of life was included (i.e. the prediction was not simply due to being a patient measure).

Design limitations include the absence of CAN patient ratings at follow-up, only measuring outcomes at two time points, not assessing patients out of contact with services at follow-up. It should of course be emphasised that these all arise from the naturalistic focus of the South Verona Outcome Project – a complete battery of assessments cannot be measured repeatedly in routine services, and the main focus of clinical services is on patients still in contact [16].

These findings could be further exploited in two ways. First, future research will need to evaluate multiple outcomes at multiple time points for all patients (whether in contact with services or not), to allow temporal prece-

dence to be established – does reducing unmet need precede quality of life improvements? The analytical strategy in such a repeated-measures design would then focus on change in needs and quality of life, which was not possible in this study due to the absence of follow-up patient CAN data. It would also allow confirmation that 'needs' and 'quality of life' are distinct constructs – which has been assumed in this study. Research which will allow these questions to be investigated are currently underway: the ongoing South Verona Outcome Project; the European Union-funded MECCA Study [17], and the Medical Research Council-funded Feedback of Outcome to Users and Staff (FOCUS) studies [18]. Second, this effect should be confirmed in randomised controlled trials which test the effect on quality of life of treatment strategies specifically designed to meet patients' needs. Such a study would also allow investigation of whether objective indicators of quality of life also change when needs are met – an important question since subjective (as measured in this study) and objective indicators of quality of life are poorly associated [9].

The cross-sectional evidence of association between patient-rated unmet need and subjective quality of life replicated previous studies investigating this question [4, 5]. Given the consistency of results from three distinct databases across two countries using both routine data in this study and research data in previous studies, the association between high patient-rated unmet need and low quality of life appears increasingly robust. Furthermore, other mental health studies indicate that no association is found when CAN data are omitted. No predictors of 6-month follow-up quality of life were found among baseline variables using 1994 data from South Verona [8], and the same psychological variables which were associated cross-sectionally were found to be the main predictors of quality of life at 2-year follow-up [9]. Unmet need also appears to have a stronger relationship with quality of life than met need – a Swedish study of 112 long-term mental patients which aggregated the number of met and unmet needs found that change in total CAN score predicted change in the health dimension of quality of life over 18 months, but not in overall quality of life [19]. This is consistent with the finding from the current study that both met and unmet need are negatively correlated with quality of life. However, the analysis also indicated other cross-sectional predictors of higher quality of life, including higher disability and a psychotic diagnosis – perhaps providing support for the 'disability paradox' in which people with disabilities report high subjective quality of life.

The graphical chain model shown in figure 1 demonstrates that there is a partition between staff and patients ratings. This is indicated by the (patient-rated) quality of life being associated only with patient-rated needs and quality of life at baseline, and staff-rated needs at follow-up only being associated with staff-rated needs at baseline. This accords with findings from previous studies that staff and patient ratings of need differ systematically [3, 20–23]. In summary, the number of needs identified by staff and patients are broadly similar, with a tendency for staff to identify slightly more needs, but the domains of need identified can differ substantially. There is more agreement between staff and patients on met needs than on unmet needs, and on domains of need with a relatively defined service response (such as Accommodation) compared with those without (such as Intimate relationships).

Clinical Implications

Clinically, this study provides the first indications that high levels of patient-rated (but not staff-rated) unmet need may, unlike symptomatology, disability or functioning, actually *cause* low levels of subjective quality of life. One explanation of this finding would be that needs are the mediating link between subjective quality of life and all its influences (rather than just psychiatric influences), since a CAN assessment identifies needs no matter what the cause. If confirmed by future research, this is an important finding – it means that if the goal of mental health services is to improve quality of life, then care should be provided differently. There should be a greater emphasis on assessing needs from the patient's viewpoint, and then identifying which of the identified unmet needs can be met. In other words, the patient's assessment of their unmet needs should more strongly inform the planning and provision of mental health care. Interventions would then be focused on meeting unmet needs, as well as on symptoms and diagnosis. This model could be tested in a randomised controlled trial comparing needs-focused with more traditional, symptom-focused mental health services.

However, the magnitude of association found in this study was relatively modest – meeting one unmet need would lead to a one percent change in subjective quality of life (with the 95% confidence interval indicating the maximum change consistent with the data as 2%). The finding that mental health care has a limited impact on quality of life is consistent with results from other areas of medicine, such as cardiac care [24]. The current exclusive focus on mental health outcomes as traditionally defined may pro-

duce statistically significant change, but is unlikely to provide enough leverage to produce clinically significant improvement in quality of life. If the purpose of care is to improve substantially the patient's quality of life, then future research will need to consider a much broader range of variables, including the patient's assessment of their symptoms [25] and, more generally, the patient's personality [9], motivation [26] and expectations [13], along with society-level factors such as culture [27] and economic prosperity [28]. Such research could have profound implications for future mental health care.

Acknowledgements

We are grateful to the patients who participated in the South-Verona Outcome Project and to the staff of the South-Verona community psychiatric service for their kind collaboration in the project. We thank Prof. Annibale Biggeri, Dr. Giulia Bisoffi and Dr. Chiara Bonetto for statistical advice, Dr. Antonio Lasalvia for data management advice, and Annalisa Voccia for secretarial assistance. The South Verona Outcome Project was funded by Fondazione Cassa di Risparmio di Verona, Vicenza, Belluno e Ancona, Progetto Sanità 1996–1997 Grant 'The Role of Social Factors on Onset, Course and Outcome of Mental Disorders. An Epidemiological Approach' to Professor Michele Tansella. The paper was written as part of a Medical Research Council Clinician Scientist Fellowship awarded to M.S. (G108/390).

References

1 Phelan M, Slade M, Thornicroft G, Dunn G, Holloway F, Wykes T, Strathdee G, Loftus L, McCrone P, Hayward P: The Camberwell Assessment of Need: The validity and reliability of an instrument to assess the needs of people with severe mental illness. Br J Psychiatry 1995;167:589–595.

2 Department of Health: National Service Framework for Mental Health. London, HMSO, 1999.

3 Lasalvia A, Ruggeri M, Mazzi MA, Dall'Agnola RB: The perception of needs for care in staff and patients in community-based mental health services. The South Verona Outcome Project 3. Acta Psychiatr Scand 2000;102:366–375.

4 UK700 Group: Predictors of quality of life in people with severe mental illness. Study methodology with baseline analysis in the UK700 trial. Br J Psychiatry 1999;175:426–432.

5 Slade M, Leese M, Taylor R, Thornicroft G: The association between needs and quality of life in an epidemiologically representative sample of people with psychosis. Acta Psychiatr Scand 1999;100:149–157.

6 Bollen K: Structural Equations with Latent Variables. New York, Wiley, 1998.

7 Ruggeri M, Gater R, Bisoffi G, Barbui C, Tansella M: Determinants of subjective quality of life in patients attending community-based mental health services. The South-Verona Outcome Project 5. Acta Psychiatr Scand 2002; 105:131–140.

8 Ruggeri M, Biggeri A, Rucci P, Tansella M: Multivariate analysis of outcome of mental health care using graphical chain models. The South-Verona Outcome Project 1. Psychol Med 1998;28:1421–1431.

9 Ruggeri M, Bisoffi G, Fontecedro L, Warner R: Subjective and objective dimensions of quality of life in psychiatric patients: A factor analytical approach: The South Verona Outcome Project 4. Br J Psychiatry 2001;178:268–275.

10 Endicott J, Spitzer RL: The Global Assessment Scale: A procedure for measuring overall severity of psychiatric disturbance. Arch Gen Psychiatry 1976:33;766–771.

11 Ventura J, Green MF, Shaner A, Liberman RP: Training and quality assurance with the Brief Psychiatric Rating Scale: 'The drift busters'. Int J Methods Psychiatr Res 1993;3:221–244.

12 World Health Organization: Disability Assessment Schedule (DAS-II). Geneva, World Health Organization, 1998.

13 Ruggeri M, Dall'Agnola R: The development and use of the Verona Expectations for Care Scale (VECS) and the Verona Service Satisfaction Scale (VSSS) for measuring expectations and satisfaction with community-based psychiatric services in patients, relatives and professionals. Psychol Med 1993;23:511–523.

14 Oliver JP: The social care directive: development of a quality of life profile for use in community services for the mentally ill. Soc Work Soc Sci Rev 1991;3:4–45.

15 Edwards D: Introduction to Graphical Modelling, ed 2. New York, Springer, 2000.

16 McKenzie N, Marks I: Routine monitoring of outcome over 11 years in a residential behavioural psychotherapy unit. Psychother Psychosom 2003;72:223–227.

17 Priebe S, McCabe R, Bullenkamp J, Hansson L, Rossler W, Torres-Gonzales F, Wiersma D: The impact of routine outcome measurement on treatment processes in community mental health care: Approach and methods of the MECCA study. Epidemiol Psychiatr Soc 2002; 11:198–205.

18 Slade M: Routine outcome assessment in mental health services. Psychol Med 2002;32: 1339–1344.

19 Björkman T, Hansson L: Predictors of improvement in quality of life of long-term mentally ill individuals receiving case management. Eur Psychiatry 2002;17:33–40.

20 Slade M, Phelan M, Thornicroft G, Parkman S: The Camberwell Assessment of Need (CAN): Comparison of assessments by staff and patients of the needs of the severely mentally ill. Soc Psychiatry Psychiatr Epidemiol 1996;31: 109–113.

21 Slade M, Phelan M, Thornicroft G: A comparison of needs assessed by staff and an epidemiologically representative sample of patients with psychosis. Psychol Med 1998;28:543–550.

22 Wiersma D, Nienhuis FJ, Giel R, Slooff C: Stability and change in needs of patients with schizophrenic disorders: A 15- and 17-year follow-up from first onset of psychosis, and a comparison between 'objective' and 'subjective' assessments of needs for care. Soc Psychiatry Psychiatr Epidemiol 1998;33:49–56.

23 Hansson L, Vinding HR, Mackeprang T, Sourander A, Werdelin G, Bengtsson-Tops A, Bjarnason O, Dybbro J, Nilsson L, Sandlund M, Sorgaard K, Middelboe T: Comparison of key worker and patient assessment of needs in schizophrenic patients living in the community: A Nordic multicentre study. Acta Psychiatr Scand 2001;103:45–51.

24 Beck CA, Joseph L, Belisle P, Pilote L: Predictors of quality of life 6 months and 1 year after acute myocardial infarction. Am Heart J 2002; 142:271–279.

25 Lasalvia A, Ruggeri M, Santolini N: Subjective quality of life: Its relationship with clinician-rated and patient-rated psychopathology. Psychother Psychosom 2002;71:275–284.

26 Grahn B, Ekdahl C, Borgquist L: Motivation as a predictor of changes in quality of life and working ability in multidisciplinary rehabilitation. Disabil Rehabil 2002;22:639–654.

27 Michalos AC, Zumbo BD: Ethnicity, modern prejudice and the quality of life. Soc Indicators Res 2001;53:189–222.

28 Schyns P: Crossnational differences in happiness: Economic and cultural factors explored. Soc Indicators Res 1998;43:3–26.

What outcomes to measure in routine mental health services, and how to assess them: a systematic review

Mike Slade

Objective: Routine outcome assessment in adult mental health services involves the ongoing assessment of patient-level outcomes. Use of outcomes to inform treatment is widely recommended, but seldom implemented. The goals of this review were (i) to identify principles that have been proposed for implementing routine outcome assessment, (ii) to identify the full range of outcome domains that have been proposed for assessment, and (iii) to synthesize proposals for specific outcome domains into emergent categories.
Method: A systematic review of published and unpublished research was undertaken, using electronic databases, research registers, conference proceedings, expert informants and the World Wide Web. For goal (i) studies were included that proposed principles for implementing routine outcome assessment. For goal (ii) studies were included that identified at least two patient-level outcome domains for patients using adult mental health services and made some reference to a broader literature base.
Results: Six thousand four hundred publications matched initial search criteria. Seven distinct sets of principles for choosing patient-level outcomes were located, which showed a fair degree of consensus. Sixteen outcome domain proposals were identified, which were synthesized into seven emergent categories: wellbeing, cognition/emotion, behaviour, physical health, interpersonal, society and services.
Conclusions: The findings from this review were used to develop a four-step method for adult mental health services wishing to implement routine outcome assessment.
Key words: mental health, outcome assessment.

Australian and New Zealand Journal of Psychiatry 2002; 36:743–753

A standard definition of 'outcome' in mental health care is 'the effect on a patient's health status attributable to an intervention by a health professional or health service' [1, p.3]. Despite being widely used, this definition has been challenged for several reasons – outcome could result from self- rather than professional help, the link between intervention and outcome is not straightforward, outcomes are not always positive, outcomes may be influenced by the absence rather than presence of an intervention, and outcome may differ from different perspectives [2]. There is as yet no consensus about an agreed definition of outcome for individual patients. Further complexity arises when evaluating outcome in mental health services. Three levels of mental health service can be identified: specific treatments, combinations of treatments (such as a community mental health centre) and population-wide treatments (all programmes for a defined population, such as a managed care organization) [3]. The outcome data needed to inform each level are very different.

Perhaps because of this complexity, the systematic measurement of outcome in mental health services has traditionally been the preserve of researchers. In general, most efforts to assess outcome take place in 'research

Mike Slade, Medical Research Council Clinician Scientist Fellow

Health Services Research Department (PO29), Institute of Psychiatry, De Crespigny Park, Denmark Hill, London, SE5 8AF, United Kingdom.
Email: m.slade@iop.kcl.ac.uk

Received 19 March 2002; 2nd revision 23 July 2002; accepted 30 July 2002.

contexts where specifically funded and trained external raters parachute into routine clinical settings in order to guarantee the validity and reliability of study measures' [4]. However, in the last decade a new international emphasis on the ongoing measurement of outcome in routine mental health services – routine outcome assessment – has begun to emerge [5–8].

Internationally, a range of approaches have been taken to assess outcome routinely. In the US, the focus on outcomes as the measure of success for mental health services has been driven by cost-containment. Difficulties in implementing Diagnostic Related Groups [9] and the growth in proportion of healthcare costs spent on mental health services from 4% in the early 1980s to 25% in the early 1990s [8] led to an emphasis on the use of outcomes. This emphasis fits with other quality enhancement initiatives, the increasing availability of an information technology infrastructure, the importing of 'learning organization' values from business into health care, and the pressure from consumer organizations for issues of choice, quality and value to be considered [8].

In the UK, a similar impetus has come from the drive towards evidence-based medicine. This is exemplified by the development of the National Institute for Clinical Excellence to provide cost-effectiveness information about health interventions, and the development of national standards for mental health care [10]. Other influences include an emphasis on clinical governance and practice guidelines, a political emphasis on quality and on patient experience, the development of high-profile (if not widely used) outcome measures for routine clinical use [11] and a societal shift towards consumerism, with concomitantly increased expectations about mental health services.

Most approaches to collecting data within routine mental health services have therefore been intended to inform programmes and systems. However, there is an increasing recognition of the importance of treatment-level outcomes, which can inform the future care provided to individual patients [2]. Australia has the most coherently developed approach to treatment-level routine outcome assessment. A systematic review of patient outcomes was undertaken as part of the first national mental health strategy, resulting in proposals for specific assessments to use routinely [1]. These assessments were then independently field-tested, to evaluate their utility [12]. The resulting recommendations have informed Australian practice in routine outcome assessment.

Mental health services implementing routine outcome assessment will want to base their efforts on principles developed through the experience of other services. One specific decision will be the outcome domains (conceptually distinct components of outcome, such as quality of life, symptomatology or satisfaction with care) to assess. This article is intended to assist adult mental health services in implementing routine outcome assessment, by using a systematic review of the available literature to inform a proposed method for implementation. The review goals were (i) to identify the principles that have been proposed for implementing routine outcome assessment, (ii) to identify the full range of outcome domains that have been proposed for assessment, and (iii) to synthesize proposals for specific outcome domains into emergent categories.

Method

Study selection

The main sources for published information were the electronic databases shown in Table 1. However, electronic searching will not identify all relevant research, partly through missing relevant indexed journal papers, and partly through not accessing technical reports, discussion papers and other forms of 'grey literature' [13]. Efforts were made to access these studies using three methods. First, researchers active in the field were consulted, and findings presented at the four European Network for Mental Health Service Evaluation conferences were reviewed. Second, the World Wide Web was searched using Copernic 2000, an internet search engine which collates the findings from other search engines. Third, the UK National Research Register and the Research Findings Electronic Register were searched. No language restrictions were employed in any search, and non-English articles were included where an abstract in English was available. Pre-publication and 'in press' manuscripts were included and the literature review was completed by the author.

It was not possible to identify a search strategy that differentiated between publications relating to mental health research and to routine mental health services. Both were therefore included. Some of the outcome domains identified were described as models of 'health status', 'wellbeing' or 'quality of life', but no search strategy was identified that allowed searching on any of these key words with sufficient specificity. Similarly, no satisfactory synonym for routine (as in 'routine mental health services') could be found, so this aspect was incorporated where possible when reviewing abstracts (although often the distinction between research and routine clinical uses was not made). To maximize sensitivity, the search strategy was deliberately over-inclusive.

Several electronic databases were searched and the *MEDLINE* search engine allowed the most comprehensive search strategy. For the *MEDLINE* search, all studies relating to mental health or psychiatry (identified from title, abstract or medical subject heading (MeSH) heading) with either the word 'outcome' in their title or abstract or a MeSH heading of 'Outcome and process assessment (mental health)' were identified. Treatment trials and animal-only studies were excluded. The search was then adapted for other electronic databases, which were less sophisticated. For instance, the IBSS engine only allowed one search term, so 'outcome' was used. Duplicates of all identified articles were removed using Reference Manager Professional Edition, Version 9.5 (ISI ResearchSoft, Berkeley, California).

Table 1. Electronic databases used for the literature review, and number of publications from each database matching initial search criteria

Name of database	Brief description	Search engine	Web site	Search dates	Number found
Primary sources					
Medline	Human medicine and biomedical research	Ovid version 7.8 accessed via Biomed	http://www.biomed.niss.ac.uk	1993 – June 2000	1973
Cumulative Index of Nursing and Allied Health Literature (CINAHL)	Nursing literature	Ovid version 7.8 accessed via Biomed	http://www.biomed.niss.ac.uk	1993 – April 2000	255
PsycINFO	Psychology research	Ovid version 7.8 accessed via PsycINFO	psycinfo.umds.ac.uk	1993 – April 2000	1941
International Bibliography of the Social Sciences (IBSS)	Social Sciences database	Ovid version 7.8 accessed via BIDS	http://www.bids.ac.uk	1951 – June 2000	539
National Research Register 2000, Issue 3 (completed projects only)	UK National Health Service	Update software	http://www.update-software.com	Unspecified	615
Research Findings electronic Register (ReFeR)	UK National Health Service	Unspecified	tap.ccta.gov.uk/doh/ refr_web.nsf/ basicsearch	1995 – August 2000	147
World Wide Web	'grey' literature	Copernic 2000	http://www.copernic.com	August 2000	450
Secondary sources					
The Cochrane Methodology Register	Articles on the science of research synthesis	Update software	http://www.update-software.com	Unspecified	161
NHS Economic Evaluation Database	Economic evaluation of health interventions	NHS CRD	http://www.york.ac.uk/Institute/crd	1994 – June 2000	248
Health Technology Assessment database	Information on health technology assessments	NHS CRD	http://www.york.ac.uk/Institute/crd	1993 – January 2000	28

Since a high-quality review was published in 1994 [1], the electronic search was restricted to publications in or since 1993. The review was undertaken between September and October 2000.

Data extraction

The inclusion criterion for principles was that (potentially) measurable principles were proposed for implementing routine outcome assessment. The main reasons for exclusion were that proposals were too narrow (e.g. relating to minimizing staff resistance to outcome measurement, relating to measuring outcome of psychotherapy, relating to outcome data solely for service funders, and (most commonly) relating to desirable psychometric properties of assessments) or too general (e.g. relating to measuring outcome in all medical settings).

The inclusion criteria for outcome domains was that the proposal identified a range of (i.e. more than one) treatment-level outcome domains for patients using adult mental health service, and made some reference to a broader literature base beyond personal experience or expertise. Proposals relating to other areas of medicine were only included if the proposal was sufficiently generic to have relevance to mental health services, as rated by the reviewer. Exclusion criteria included a focus on choosing outcome measures (rather than domains), proposals relating to patient groups other than adults of working age or outcome domains for programmes (e.g. service uptake) or systems (e.g. interagency working) and reports of individual treatment trials. Emergent categories of outcome domains were identified, comprising the smallest conceptually distinct set of categories which could encompass all proposed outcome domains.

The titles of all publications identified in the initial electronic search were read, to identify those with possible relevance. The abstracts from these identified publications were then reviewed, and where they appeared to meet inclusion criteria the full publication was obtained and read, following which a decision was made as to its inclusion. The reference lists from all obtained articles were also hand-searched for relevant earlier publications. Where more than one publication referred to the same piece of work, only the earliest was included, even where the apparently later one indicated that it was the first publication (e.g [14,15]). Where the date of 'publication' for grey literature was not clear, the date of the latest citation was used (e.g [16]).

Results

The databases searched are shown in Table 1, with the number of publications matching search criteria shown in the final column.

As well as the 6357 publications identified electronically, approximately 50 were identified from nonelectronic sources. Approximately 150 full papers were obtained. Seven papers presenting distinct sets of principles for implementing routine outcome assessment were located [5,12,14,17–20], which together identified 18 principles. Table 2 orders these principles by degree of consensus (with one principle identified in 6 of the 7 studies, two identified in 5 studies, etc.).

The authors of six of the studies worked in North American institutions and one [12] in Australia. Five studies were conducted under the auspices of national bodies – the National Institute for Mental Health [17,18] and the National Alliance for the Mentally Ill [14], university departments [20] and Government departments [12], and two by individuals [4,19]. Studies [14] and [17] are based on the findings of task

Table 2. Principles identified in 7 studies for implementing routine outcome assessment in mental health services

Principle	Reference																				
	5	17	18	19	20	14	12														
Standardized measures should be used	✓	✓																			
Relevance to informing practice should be emphasized		✓																			
Multiple perspectives should be used		✓																			
Standardized methods should be used		✓																			
Data collection should be cheap and simple		✓																			
Measures should be relevant to the patient group		✓																			
Treatment received should be characterized	✓																				
Feedback sould be quick, easy and meaningful	✓	✓																			
Aggregated data should be comparable with benchmarks	✓	✓																			
Meaning of measures should be comprehensible		✓																			
Data should be collected longitudinally																					
Casemix (e.g. diagnosis) should be assessed																					
Measures should show means/processes of change		✓																			
Measures should fit with psychopathology theories		✓																			
Outcomes chosen should be multidimensional																					
Costs should be included																					
Data on treatment leavers should be collected†	✓																				
Individual utility differences should be considered	✓																				

†The outcome domains chosen are specific to the individual, since different patients may attribute different degrees of importance to particular outcome domains.

forces, studies [18] and [20] on expert panels, study [12] on a literature review (strongly based on [17] and [21]), and studies [9] and [19] on personal expertise.

Sixteen specific and distinct proposals for outcome domains to measure were identified. These are shown chronologically in Figure 1, organized using seven categories that emerged from the review process. The original terms for proposed outcome domains are shown, with vertical lines indicating the span of the outcome domain across the emergent categories.

All literature reviews used as justification for the proposed domains were selective – no systematic review was identified. Study [34] stood out as being underpinned by a sound (though not systematic) literature review. It included aspects such as sentinel events (undesirable outcomes of a magnitude to always warrant a detailed investigation of a clinician's actions) and technical proficiency (of the clinician), which featured prominently in the general medical but not the mental health literature. (Technical proficiency is regarded as a process rather than an outcome domain in most mental health-focused literature).

The affiliations of the authors of studies [13,19–23,25–28,31] were North American, of study [24] was Italian, of studies [29,30] were English, and of studies [32,33] were English and Italian. Studies [19,20] described domains of health-related quality of life, studies [21,28,30] described domains of health status, and the remainder described treatment outcome domains. Study [19] related to a cardiovascular patient group, studies [20,21,31] to general medicine, [28] to psychotherapy, [30] to clinical trials, and the remainder to patients of mental health services.

Most proposals defined the meaning of the outcome domain. For example, Ware [25] defined 'mental health' as both behavioural dysfunction and the frequency and intensity of symptoms of psychological distress and feelings of psychological wellbeing; 'physical health' as limitations in performance, ability to perform daily self-care, or undertake a range of physical activities; 'social functioning' as both social contacts and social ties or resources; and 'role functioning' as performance of role activities such as employment, school and housework. Hargreaves and Shumway [24] stated that humanistic goals are to maximize the patient's and the family member's sense of wellbeing and personal fulfillment; clinical goals are to improve or cure an illness or disorder, reducing or eliminating its signs and symptoms; rehabilitative goals are to restore or improve social and vocational functioning; and public safety goals are to prevent injury whether from assaultive or self-destructive behaviours that arise out of illness, or from 'destructive' (i.e. iatrogenic) side-effects of the services themselves. Finally, Campbell [16] described wellbeing as linked to the protection of a person's basic human freedoms, safety and privacy; personhood as a recognition of common humanity and a tolerance for individual differences; self-help as including both self-help groups and provision of specific services by consumers; recovery as the maximization of a consumer's life and the minimization of their illness with appropriate, relevant and continuous flexible service and supports collaboratively developed and chosen; empowerment as involving the help receiver having direct control over the help and there being reciprocity between help givers and receivers; iatrogenic effects and negative outcomes being undesired consequences from or side-effects of receiving certain public mental health services or treatments; and satisfaction and dissatisfaction (both being important to measure) relating to the consumer's view of services received and the results of the treatment.

Study	Wellbeing	Cognition/ emotion	Behaviour	Physical health	Interpersonal	Society	Services
Wenger [22]	Perceptions of health status & wellbeing	Intellectual Emotional Symptoms	Daily routine	Symptoms – other illnesses	Social	Economic	
Bergner [23]	Health perceptions General life satisfaction	Symptoms Emotional status Cognition	Functional status	Sleep and rest Energy and vitality	Role activities Social functioning		
Hargreaves [24]	Humanistic	Clinical			Rehabilitative	Public safety	
Ware [25]	Mental health			Physical health	Social functioning Role functioning		
Rosenblatt [26]	Life satisfaction and fulfillment	Clinical			Functional	Welfare and safety	
Ruggeri [27]	Quality of life Needs for care	Psychopathology			Social functioning and support	Burden of relatives	Satisfaction
Sederer [28]		Symptom			Functional		Satisfaction with treatment
Cook [29]	Quality of life				Vocational Educational Residential		Hospitalization Consumer satisfaction
McGlynn [30]	Quality of life	Clinical			Functional		Adverse events Satisfaction with medical care
Schloster [31]	Life satisfaction Life direction	Emotional Mental		Physical		Social	
Campbell [16]	Wellbeing Personhood	Recovery					Self-help Empowerment Iatrogenic effects and negative outcomes Satisfaction and Dissatisfaction
Clifford [32]		Psychological	Activities of daily living	Physical wellbeing	Interpersonal relationships	Social circum-stances	Response to care
Fitzpatrick [33]	Psychological wellbeing Personal constructs	Cognitive functioning		Physical function	Social well-being Role activities		Satisfaction with care
Jennings [34]	Health status/ health-related quality of life Patient knowledge	Diagnosis Psychological function Symptom management	Behavioural Activities of daily living	Comfort/dis-comfort Physical function Mobility Disability	Social function		Patient satisfaction Appropriateness of treatment Sentinel events Technical proficiency
Thornicroft [35]	Quality of life	Disabilities Needs				Impact of caring	Satisfaction with services
		Symptom severity					
Thornicroft, 2000 [36]	Quality of life	Needs				Carer burden	Quality of care Satisfaction
			Global functioning		Social disabilities		

Figure 1. Outcome domain proposals from 16 studies for use in mental health services, grouped into 7 emergent categories.

Discussion

This study reviewed principles and outcome domains for routine outcome assessment in adult mental health services. Some agreement exists about the key principles for routine outcome assessment. There is consensus on the need for scientific rigour in routinely collected data to be achieved through the research strategies of standardized and relevant measures and methods, assessment from multiple perspectives, and that putting outcome information into context requires information on treatment received. There is also consensus about the need to ensure that the collection of data requires minimal effort and provides relevant information. There is less consensus about the specifics, since routine outcome assessment can be done for different purposes. Where the goal is to provide outcomes to inform the comprehensive provision of mental health care for a defined population, there is a need for benchmarking and cost information. Where the goal is to inform the planning, development and evaluation of a specific service (such as a community mental health team), the focus is more on casemix and treatment leavers. Finally, at the treatment level it is important to collect data longitudinally, and to consider the preferences (utility) of individual patients.

Two themes emerged from the review of outcome domains. First, early proposals did not include assessment of the experience of receiving services, which only came to prominence in the mid-1990s. Second, two distinct perspectives can be identified. Publications within the medical literature used a more psychiatric language, emphasized the staff perspective, and had a focus on the amelioration of disability. The only publication that was located from outside the medical literature [16] used a more phenomenological language emphasizing the patient's experience of care, and focused on increasing the patient's wellbeing and avoiding harm from and dependence on mental health services. The search strategy was systematic within the psychiatric literature but not within the broader social science or user movement literature. Therefore other well-developed proposals for outcome domains probably exist that were not identified in this review.

Synthesizing previous work led to the emergence of seven categories of outcome domains: wellbeing, cognition/emotion, behaviour, physical health, interpersonal, society and services. Wellbeing relates to the patient's sense of subjective wellbeing in their life (not about services), and by definition can only be assessed by the patient. This may involve considering individual life domains, or be a single global outcome. The next three categories relate to the patient as an individual – their cognition/emotion, their behaviour and their physical health. For all three of these the clinician and the patient may have their own assessment, and their assessments may differ. For example, in the cognition/emotion category the self-reported level of depression may not accord with the clinician's assessment of 'objective' signs of depression. The interpersonal category refers to aspects of the patient in relationship to others, both in individual social interactions and in performance of social roles. The society category describes aspects of a patient's mental health problems that may impinge on wider society, both at the individual level of the burden of care, and the macro-level of costs (e.g. welfare benefits, reduced public safety). Finally, the services category emerged as a distinct outcome domain to consider, including both positive and negative aspects of receiving mental health care.

Internal validity

The internal validity of this review can be considered in terms of the criteria outlined in the Quality of Reporting of Meta-analyses guidelines [37]. The review could be improved in a number of ways. The inclusion criteria could not be formally specified beyond the conceptual level, since relevant principles and outcome domain proposals appeared in different contexts. This difficulty in constructing a precise search strategy for a non-quantitative search has been acknowledged by systematic review specialists: 'when searching for qualitative research for the purpose of systematic reviews, it is often not practicable to construct strategies to capture the many ways in which such research may be described [38].' Similarly, characterizing the identified studies was problematic. The intended type of study was clearly described – related to routine outcome assessment in adult mental health services. In practice, included publications often were not clear about their remit, and hence were difficult to characterize. No flow profile was maintained, to show the points of and reasons for attrition. Approximately 6400 publications matched initial search criteria, but no record was then kept of numbers excluded at each stage (e.g. removal of duplicates, initial screening, removal following retrieval). The rationale for this was that such information would not be relevant to a non-quantitative review, although as the review progressed it became clear that some synthesis was possible, and that the attrition rate would in any case have been of interest for identifying the key points of exclusion. These methodological deficits reduce the replicability of the review.

The quality of the research was not assessed. Some commentators suggest that no ranking of qualitative methods is possible and each article needs to be considered

on its own merits (e.g [39,40]). Others have developed approaches to judging the quality of quantitative research (e.g [41,42]), although these relate in the main to methodological standards rather than the evaluation of conceptual work. The requirement for some form of literature review for outcome domains was a minimum quality assurance approach, but this could be strengthened, for example by duplicate reviewing or the involvement of more than one reviewer.

Generalizability

Is the review externally valid? Eleven (73%) of the sets of principles and six (86%) of the outcome domain proposals come from North American authors, reflecting that purchaser-driven pressures have stimulated more activity in routine outcome assessment there than anywhere else. Report cards and other means of characterizing aspects of the effectiveness of a service are now routinely used in North America, which has had a substantial impact on the types of care available, and the length of time for which it can be offered. The findings of this review are therefore of most relevance to North American settings. However, the Australian, UK and Italian studies were compatible with the North American work, and hence the findings are likely to have external validity more broadly.

Clearly conceptualization of health and illness differs across cultures, with some cultures giving prominence to domains (such as culture or spirituality) that were not located in this study. Some population subgroups are insufficiently researched to allow the identification of valid understandings of what constitutes normal and abnormal within the culture, such as the paucity of knowledge concerning mental health for Aboriginal youth [43]. Therefore, the values and aspirations of individual subgroups (e.g. Maori [44]) have probably not been captured in the outcome domains identified in this study.

Overall, the external validity of this review for different groups is difficult to establish without a broad range of empirical data. However, there is no reason to think that using the seven emergent categories of outcome domains as a starting point for implementing routine outcome assessment would be unwise.

Clinical implications

What are the clinical implications of this review? It provides a starting point for mental health services that are considering the use of patient-level routine outcome assessment. Such services might work towards decisions about methods and measures in four steps.

Step 1 – realism

Consider the principles for routine outcome assessment identified in Table 2. They are useful to consider for two reasons. First, they give an indication of the range of issues that will need to be addressed. The psychometric properties of any outcome measures should be established – locally developed assessments are unlikely to be appropriate – and they will need to be administered in a standard way. As a minimum, both patients and staff will be used as informants, and all collected data must either inform the treatment of individual patients or the development and evaluation of services. To make sense of the data, some characterization of the treatment received will be necessary, and mechanisms for analysing the data and producing feedback will be required.

Second, consideration of these issues gives an indication of the resources required. Resources include leadership, expertise, support staff, information technology (e.g. access to computers, easy-to-use software) and clinical time. If these resources are not available then routine outcome assessment should not be undertaken. Starting to use outcome measures without the requisite resources typically results in the haphazard collection of low-quality data that is not analysed or used, until the endeavour is either abandoned or covertly sabotaged (e.g. by 100% non-response rate). Such effort is of no benefit to patients, and creates an unhelpful belief for clinicians that outcome assessment is a clinical burden rather than providing useful data to inform treatment planning.

Clinical time – a key resource

A particularly valuable resource is the clinician's time – is it best spent assessing outcome? Completing and analysing simple outcome measures in routine clinical practice can add 10% to the time spent by the clinician per patient [45]. At present, outcome measures are not used routinely within mental health care [46], suggesting that clinicians remain unconvinced that this extra time (and the consequent reduction in number of patients they can see) is a price worth paying. Indeed, since seeing patients is often viewed as valuable clinical activity and filling in forms is not seen as 'work', there are in fact active disincentives to staff completing outcome measures [47]. To make routine outcome assessment more realistic, developments may be necessary in the culture of clinical practice, the research base, and the implementation strategies.

A change in the culture of clinical practice may be needed, in which structures (e.g. number of beds) and

processes (e.g. number of clinical contacts) are de-emphasized, and outcomes become the central influence on decision-making about continuing, changing or ending care. This shift would of course have a profound impact on the way mental health services are formed, operated and evaluated. For example, demand for mental health services has increased by a factor of 4.5 from 1971 to 1997 [48], and presumably will continue to rise. If mental health services operate with an increasing and effectively unlimited caseload size, then it is unrealistic to expect any intervention requiring more time to be spent per patient to be implemented, whatever its merits. Alternative models, (e.g. limiting caseload sizes to ensure a defined level of quality of care is possible), might need to be in place before routine outcome assessment could be realistically considered. The resulting population-level health gain from these and other models of service provision could be investigated, and this type of mental health services research is urgently needed. Other examples of approaches to changing the working culture include the introduction of payment incentives for clinicians who collect and use outcome information [47], and the monitoring of outcome data during clinical training [49,50].

Research studies are needed that quantify the effectiveness and cost-effectiveness of implementing routine outcome assessment. This will allow informed discussion about the relative merits of different styles of clinical practice, such as providing care explicitly targeted at improving outcomes versus providing care audited for its conformance to good practice. Several such studies are currently underway across Europe [e.g. 51,52] that will provide this evidence.

Finally, creative approaches will need to be developed that minimize the time spent by clinical staff in collecting and analysing outcome data. For instance, some uses may only require patient-rated data. In these cases, the use of electronic questionnaires for data collection could be considered. For analysis, the use of automatic data entry and analysis or the employment and training of non-clinical staff for this purpose could be evaluated. One approach would be the provision of a computer for a patient to use before his/her clinical meeting, which then analyses and prints out the resulting outcome data (e.g. by using previously entered information to chart progress over time) for the clinician to review at the meeting. Such an approach would raise further questions, such as whether comparison between clinicians or between clinics is appropriate, and how to maintain data quality where the assessment is not undertaken by a clinician. The costs would include the setting up and maintenance of the computer and the clinician's time spent reading the results, and the potential benefits would be reduced assessment time and more informing of care planning [6,53–55]. Overall, minimizing the burden and maximizing the potential benefits from using outcomes in clinical care will make routine outcome assessment more realistic.

Step 2 – outcome domains

Identify what outcome domains are appropriate to monitor in the service. This decision will be informed by considering the seven emergent categories identified in Figure 1. Differing understandings of mental health problems lead to disagreement about the outcome domains to consider. These understandings can be conceptualized as lying on a spectrum, from patient-defined to professionally defined [56]. At the extreme of the patient-defined end lies an understanding that emphasizes the importance and uniqueness of individual experience, and accords no value to comparison of one person with another. At the extreme of the professionally defined end lies an understanding that emphasizes the importance of using scientific knowledge to understand abnormal mental experiences, and accords no value to the meaning attached to these experiences by the patient. Most clinical practice, of course, takes place within these extremes, but the point on the continuum will influence the outcome domains selected for routine assessment. A service operating towards the patient-defined end will be more interested in outcome domains related to individual phenomenological change, and how the health care service is experienced. A service operating towards the professionally defined end will be more interested in outcome domains related to symptoms and functioning, and ensuring that the interventions given are in accord with research evidence. One approach to reconciling the conflicting interests of staff and patients is to assess all seven domains identified in this review, which runs the risk of being impractical. Another approach is to identify one domain for assessment, which runs the risk of being insufficiently meaningful. Most approaches to implementing routine outcome assessment involve assessing between two and four domains.

Step 3 – technicalities

Consider how these outcome domains will be measured within the service. This will require consideration of several questions [33,57]. What constitutes a good outcome for a patient who is not expected to recover? Is the goal to show that the treatment caused improvement, or just to show that improvement occurred (without reference to treatment)? Whose outcome is being considered? Cost containment, adherence to clinical protocols,

reduced symptomatology and reduced visibility of the mentally ill are all outcomes from different perspectives. Is the focus just on outcome for the patient, or also for their relatives or carers in their own right? Is the focus on individual change, or aggregating data to investigate changes in groups? The data required for individual or group-level analysis may be very different. Are direct measures (e.g. from the patient or carer) or indirect measures (e.g. from staff or service usage) to be favoured? Is equal weighting given to externally observable measures and private, non-observable experiences of the patient? Are global (single-score) or multiple item measures preferred? Are generic measures (applicable to broad groups) or specific measures (for highly characterized conditions) preferred? Are individualized (tailored to the individual) or standardized measures (which can be compared to group norms) preferred? Should assessment be undertaken at 'important' times during the patient's pathway through care, or at predetermined fixed time periods? How should 'conflicting' changes, such as increased symptoms accompanied by increased quality of life, be interpreted? The answers to these questions will reflect underlying principles of the service.

Step 4 – outcome measures

Identify the outcome measure(s) that most meet the requirements that have been identified in Steps 1–3. This should be the final step. Several collations of outcome measures exist (e.g. [36,58–61]), although these tend to evaluate their suitability for use in research studies. Measures for routine use also need to be 'feasible' [1], for example by being brief, simple, acceptable, available, relevant and valuable [47]. It is possible to evaluate the feasibility of outcome measures [62].

By way of example, an evaluation of routine outcome assessment in adult mental health services is currently taking place in London. Patients with any diagnosis are included providing they are aged 18–65 and have been in contact with the mental health service for at least 3 months. The optimal frequency and format of assessment and feedback for different patient populations remains an open question, but in this study both staff and patients are asked to complete outcome measures every month, and treatment-level graphical feedback is provided to both people every three months. The goal of the intervention is to beneficially influence the process and content of care [2]. Six criteria were used for choosing outcome measures: (i) the measure either assesses a desired outcome (needs, quality of life) or process measure (therapeutic alliance), allows explicit comparison between staff and patient views, or is a severity measure leading to a desirable focus on outcome; (ii) the measure has peer-reviewed published evidence of acceptable psychometric properties; (iii) the measure is designed specifically for a mental health population; (iv) the measure is brief to administer (arbitrarily chosen as an administration time of less than 15 min); (v) there is no charge to use the measure; and (vi) there is no training required to use the measure. Using these criteria, the measures chosen for staff and patient were the Helping Alliance Scale (HAS) to assess therapeutic alliance [63], the Camberwell Assessment of Need Short Appraisal Schedule (CANSAS) to assess needs [64] and the Threshold Assessment Grid to assess severity [65]. The patient measures chosen were the HAS, the CANSAS and the Manchester Short Assessment to assess quality of life [66].

Routine assessment of outcome in mental health services can indirectly benefit patients, by informing the development of programmes and systems [8]. It also has the potential to provide valuable treatment-level information which directly benefits patients [45]. The long-term goal is for routine outcome assessment to become an integral component of clinical care, rather than an administrative burden added on to the 'real' work of clinicians [5]. Carefully thought-out and well-resourced approaches to collecting and using outcome information are therefore needed, to avoid wasting effort and clinical goodwill. This review provides a method for service managers and clinicians who want to assess the impact of care on people using routine adult mental health services.

Acknowledgements

Thanks to Graham Thornicroft and Elizabeth Kuipers for their valuable advice. This work was supported by a Medical Research Council Clinician Scientist Fellowship (G108/390).

References

1. Andrews G, Peters L, Teesson M. Measurement of consumer outcome in mental health. In: CRUFAD, eds. *Clinical research unit for anxiety disorders. A report to the National Mental Health Information Strategy Committee.* Sydney: CRUFAD, 1994.
2. Slade M. Routine outcome assessment in mental health services. *Psychological Medicine* (in press).
3. Burns T, Priebe S. Mental health care systems and their characteristics: a proposal. *Acta Psychiatrica Scandinavica* 1996; 94:381–385.
4. Harrison G, Eaton W. From research world to real world. *Current Opinion in Psychiatry* 1999; 12:187–189.
5. Ellwood P. Outcomes management – a technology of patient experience. *New England Journal of Medicine* 1988; 318:1549–1556.

6. McKenzie N, Blanes T, Marks IM. Computerised clinical benefit-cost audit of mental health care II: time input, costs, patient satisfaction. *Journal of Mental Health* 1995; 1:71–78.
7. Walter G, Kirkby KC, Marks IM, Whiteford H, Andrews G, Swinson R. Outcome measurement – sharing experiences in Australia. *Australasian Psychiatry* 1996; 4:316–318.
8. Lyons JS, Howard KI, O'Mahoney MT, Lish JD. *The measurement and management of clinical outcomes in mental health*. New York: Wiley, 1997.
9. Dickey B, Azeni H. Data watch: impact of managed care on mental health services. *Health Affairs* 1992; 11:197–204.
10. Department of Health. *Mental health national service framework*. London: HMSO, 1999.
11. Wing JK, Beevor AS, Curtis RH, Park SB, Hadden S, Burns A. Health of the Nation Outcome Scales (HoNOS). *British Journal of Psychiatry* 1998; 172:11–18.
12. Stedman T, Yellowlees P, Mellsop G, Clarke R, Drake S. *Measuring consumer outcomes in mental health.* Canberra: Department of Health and Family Services, 1997.
13. NHS Centre for Reviews and Dissemination. *Undertaking systematic reviews of research on effectiveness: CRD's guidance for carrying out or commissioning reviews*, 2nd edn. York: University of York, 2001.
14. Shern DL, Flynn LM. Outcomes roundtable: developing, evaluating and disseminating outcomes monitoring. *Behavioural Healthcare Tomorrow* 1996; 5:25–30.
15. Smith GR, Manderscheid R, Flynn LM, Steinwachs DM. Principles for assessment of patient outcomes in mental health care. *Psychiatric Services* 1997; 48:1033–1036.
16. Campbell J. *Consumerism, outcomes and satisfaction: a review of the literature.* [Cited 16 August 2002.] Available from URL: http://www.networksplus.net/fhp/madnation/citations/consumerism.htm
17. Ciarlo JA, Edwards DW, Kiresuk TJ, Newman FL, Brown TR. *The assessment of client/patient outcomes techniques for use in mental health programs.* Washington, DC: National Institute of Mental Health, 1981.
18. Attkisson C, Cook J, Karno M *et al.* Clinical services research. *Schizophrenia Bulletin* 1992; 18:561–626.
19. Burnam MA. Measuring outcomes of care for substance use and mental disorders. *New Directions in Mental Health Services* 1996; 71:3–17.
20. Smith GR, Rost KM, Fischer EP, Burnam MA, Burns BJ. Assessing the effectiveness of mental health care in routine clinical practice. *Evaluation and the Helping Professions* 1997; 20:65–80.
21. Green RS, Gracely E. Selecting a rating scale for evaluating services to the chronically mentally ill. *Community Mental Health Journal* 1987; 23:91–102.
22. Wenger N, Mattson M, Furberg C. *Assessment of quality of life in clinical trials of cardiovascular therapies.* New York: LeJacq, 1984.
23. Bergner M, Rothman M. Health status measures: an overview and guide for selection. *American Review of Public Health* 1987; 8:191–210.
24. Hargreaves W, Shumway M. Effectiveness of mental health services for the severely mentally ill. In: Taube C, Mechanic D, Hohmann A, eds. *The future of mental health services research.* Washington, DC: US Government Printing Office, 1989; 253–284.
25. Ware J. Measuring health and functional status in mental health services research. In: Taube C, Mechanic D, Hohmann A, eds. *The future of mental health services research.* Washington, DC: US Government Printing Office, 1989; 289–302.
26. Rosenblatt A, Attkisson C. Assessing outcomes for sufferers of severe mental disorder: a conceptual framework and review. *Evaluation and Program Planning* 1993; 16:347–363.
27. Ruggeri M, Tansella M. Evaluating outcome in mental health care. *Current Opinion in Psychiatry* 1995; 8:116–121.
28. Sederer LI, Hermann MD, Dickey B. The imperative of outcome assessment in psychiatry. *American Journal of Medical Quality* 1995; 10:127–132.
29. Cook J, Jonikas J. Outcomes of psychiatric rehabilitation. *New Directions in Mental Health Services* 1996; 71:33–47.
30. McGlynn E. Setting the context for measuring patient outcomes. *New Directions in Mental Health Services* 1996; 71:19–32.
31. Schlosser B. New perspectives on outcomes assessment: the philosophy and application of the subjective health process model. *Psychotherapy* 1996; 33:284–304.
32. Clifford P. M is for outcome: the CORE outcomes initiative. *Journal of Mental Health* 1998; 7:19–24.
33. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment* 1998; 2:14.
34. Jennings BM, Staggers N, Brosch LR. A classification scheme for outcome indicators. *Image Journal of Nursing Scholarship* 1999; 31:381–388.
35. Thornicroft G, Tansella M. *The mental health matrix.* Cambridge: Cambridge University Press, 1999.
36. Thornicroft G, Tansella M, eds. *Mental health outcome measures*, 2nd edn. London: Gaskell, 2000.
37. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *The Lancet* 1999; 354:1896–1900.
38. Khan KS, ter Riet G, Popay J, Nixon J, Kleijnen J. Phase 5: study quality assessment. In: Khan KS, ter Riet G, Glanville J, Sowden A, Kleinen J, eds. *NHS Centre for Reviews and Dissemination Report No. 4. Undertaking systematic reviews of research on effectiveness*, 2nd edn. York: University of York, NHS CRD, 2000; 37–50.
39. Howe K, Eisenhart M. Standards for qualitative (and quantitative) research: a prolegomenon. *Educational Researcher* 1990; 19:2–9.
40. Buchanan DR. An uneasy alliance: combining qualitative and quantitative research methods. *Health Education Quarterly* 1992; Spring:117–135.
41. Oakley A. *Experiments in knowing. gender and method in social sciences.* Cambridge: Polity, 2000.
42. Seale S, Silverman D. Ensuring rigour in qualitative research. *European Journal of Public Health* 1997; 7:379–384.
43. Hunter E. 'Freedom's just another word': Aboriginal youth and mental health. *Australian and New Zealand Journal of Psychiatry* 1995; 29:374–384.
44. Durie M. Mental health and Maori development. *Australian and New Zealand Journal of Psychiatry* 1999; 33:5–12.
45. Marks I. Overcoming obstacles to routine outcome measurement. *British Journal of Psychiatry* 1998; 173:281–286.
46. Gilbody SM, House AO, Sheldon TA. Psychiatrists in the UK do not use outcome measures: national survey. *British Journal of Psychiatry* 2002; 180:101–103.
47. Slade M, Thornicroft G, Glover G. The feasibility of routine outcome measures in mental health. *Social Psychiatry and Psychiatric Epidemiology* 1999; 34:243–249.
48. Verhaak PF, Van de Lisdonk EH, Bor JH, Hutschemaekers GJ. GPs' referral to mental health care during the past 25 years. *British Journal of General Practice* 2000; 50:307–308.
49. McDonald R, Blizard R. Quality assurance of outcome in mental health care: a model for routine use in clinical settings. *Health Trends* 1988; 20:111–114.
50. Duggan C, Marks I, Richards D. Clinical audit of behaviour therapy training of nurses. *Health Trends* 1993; 25:25–30.

51. Slade M. The use of patient-level outcomes to inform treatment. *Epidemiologia e Psichiatria Sociale* 2002; 11:20–27.
52. Ruggeri M, Biggeri A, Rucci P, Tansella M. Multivariate analysis of outcome of mental health care using graphical chain models. The South-Verona outcome project 1. *Psychological Medicine* 1998; 28:1421–1431.
53. Marks I. Rapid audit of clinical outcome and cost by computer. *Australian and New Zealand Journal of Psychiatry* 1995; 29:32–37.
54. Marks IM, Blanes T, McKenzie N. Computerised clinical benefit-cost audit of mental health care I: theoretical and practical issues. *Journal of Mental Health* 1995; 4:63–70.
55. Marks I. A computerized system to measure treatment outcomes and cost. *Psychiatric Services* 1996; 47:811–812.
56. Long A, Dixon P. Monitoring outcomes in routine practice: defining appropriate measurement criteria. *Journal of Evaluation in Clinical Practice* 1996; 2:71–78.
57. Trauer T. Issues in the assessment of outcome in mental health. *Australian and New Zealand Journal of Psychiatry* 1998; 32:337–343.
58. Barry MM, Zissi A. Quality of life as an outcome measure in evaluating mental health services: a review of the empirical evidence. *Social Psychiatry and Psychiatric Epidemiology* 1997; 32:38–47.
59. Dickerson E. Assessing clinical outcomes: the community functioning of persons with serious mental illness. *Psychiatric Services* 1997; 48:897–902.
60. Bartlett CJ, Coles E. Psychological health and well-being: why and how should public health specialists measure it? Part 2: stress, subjective well-being and overall conclusions. *Journal of Public Health Medicine* 1998; 20:288–294.
61. American Psychiatric Association. *Handbook of psychiatric measures.* Washington, DC: American Psychiatric Association, 2000
62. Slade M, Cahill S, Kelsey W, Powell R, Strathdee G, Valiakalayil A. Threshold 3: the feasibility of the Threshold Assessment Grid (TAG) for routine assessment of the severity of mental health problems. *Social Psychiatry and Psychiatric Epidemiology* 2001; 36:516–521.
63. Priebe S, Gruyters T. The role of the helping alliance in psychiatric community care: a prospective study. *Journal of Nervous and Mental Diseases* 1993; 181:552–557.
64. Phelan M, Slade M, Thornicroft G *et al.* The Camberwell Assessment of Need: the validity and reliability of an instrument to assess the needs of people with severe mental illness. *British Journal of Psychiatry* 1995; 167:589–595.
65. Slade M, Powell R, Rosen A, Strathdee G. Threshold Assessment Grid (TAG): the development of a valid and brief scale to assess the severity of mental illness. *Social Psychiatry and Psychiatric Epidemiology* 2000; 35:78–85.
66. Priebe S, Huxley P, Knight S, Evans S. Application and results of the Manchester Short Assessment of quality of life. *International Journal of Social Psychiatry* 1999; 45:7–12.

THE USE OF PATIENT-LEVEL OUTCOMES TO INFORM TREATMENT

MIKE SLADE

Estratto da

EPIDEMIOLOGIA E PSICHIATRIA SOCIALE

Volume 11, numero 1, gennaio - marzo 2002

Il Pensiero Scientifico Editore

Invited Paper

The use of patient-level outcomes to inform treatment

MIKE SLADE

Health Services Research Department, Institute of Psychiatry, London, United Kingdom

RIASSUNTO. **Scopo** – La valutazione dell'esito, qualunque sia lo scopo della valutazione, non è effettuata di routine nei Servizi di Salute Mentale europei. Questo lavoro discute i vantaggi dell'uso di dati sull'esito per la pianificazione dell'assistenza, nel settore della salute mentale, dedicata al singolo paziente e fornisce suggerimenti pratici per implementare questo nuovo modo di lavorare. **Metodo** – Viene effettuata una breve revisione della letteratura sull'utilizzazione delle misure di esito nell'America del Nord ed in Europa. Viene proposta una base concettuale per la valutazione di routine dell'esito - le misure ed i dati sull'esito da utilizzare per decidere se continuare, modificare o interrompere un trattamento. Viene sviluppato un modello psicologico cognitivo che indica che l'uso di routine delle misure d'esito migliora l'assistenza psichiatrica. Vengono quindi discussi i problemi percepiti relativi alle valutazioni di routine dell'esito e vengono identificati i principi utili alla implementazione. **Risultati** – Le misure di esito, in Nord America, vengono utilizzate soprattutto per generare dati al livello dei servizi (piuttosto che a livello del singolo paziente), mentre in Europa sono utilizzate raramente. L'utilizzazione routinaria dei dati sull'esito può invece facilitare le riflessioni sulla pratica clinica, un modello per prendere le decisioni che favorisce una qualità dell'assistenza clinica migliore rispetto al modello delle decisioni "automatiche". Un altro aspetto è quello della utilizzazione di misure di esito standardizzate disegnate a scopi di ricerca, da condurre in setting clinici. Questo aspetto è trattato mediante lo sviluppo di una nuova generazione di misure di esito esplicitamente previste per l'uso in clinica. Tuttavia, la maggior parte dei clinici tuttora non è convinta dei vantaggi delle misurazione routinarie dell'esito. Ricerche al riguardo, disegnate per affrontare questo problema, sono attualmente in corso in Europa. Per quanto riguarda l'implementazione di queste misure, il presente articolo identifica i principi scientifici necessari per massimizzarne la qualità ed i principi pragmatici necessari per massimizzarne le possibilità di successo. **Conclusioni** – L'uso di routine delle misure dell'esito si diffonderà sempre più nei Servizi di Salute Mentale europei. Ciò fornisce ai clinici l'opportunità di migliorare la qualità delle cure offerte ai pazienti.

PAROLE CHIAVE: misurazione dell'esito, servizi di salute mentale.

SUMMARY. **Objective** – The assessment of outcome for any purpose is not undertaken routinely in European mental health services. This paper discusses the merits of using outcome data to inform the planning of mental health care for individual patients, and provides practical advice to support the implementation of this new approach to working. **Method** – The use of outcomes in North America and Europe is briefly reviewed. A conceptual basis is proposed for routine outcome assessment – the ongoing measurement and use of outcome data to inform decisions about whether to continue, change or curtail treatment. A cognitive psychology model is developed which indicates that the routine use of outcomes will improve mental health care. Perceived problems with routine outcome assessment are discussed, and principles for implementation are identified. **Results** – Outcomes are used mainly for generating local-level (rather than patient-level) data in North America, and rarely used in Europe. The use of outcome data routinely may facilitate reflective clinical practice, a model of decision-making which leads to a higher quality of clinical care than automated problem-solving. One issue relates to the use of standardised assessments designed for research purposes in clinical settings, and this is being addressed through the development of a new generation of outcome measures which are explicitly designed for clinical use. However, most clinicians remain unconvinced of the benefits

Indirizzo per la corrispondenza: Dr M. Slade, MRC Clinician Scientist Fellow, Health Services Research Department, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF (UK).

Fax: +44 (0) 20 7277 1462

E-mail: m.slade@iop.kcl.ac.uk

of routine outcome assessment, and relevant research is currently underway across Europe which will address this concern. Scientific principles to maximise quality and pragmatic principles to maximise the chances of successful implementation are identified. **Conclusions** – The routine use of outcomes will become increasingly prominent in European mental health services. This provides clinicians with an opportunity to improve the quality of clinical care offered to patients.

KEY WORDS: outcome assessment, mental health services.

Ricevuto il 14.01.2002 – Accettato il 16.02.2002.

MENTAL HEALTH SERVICES

Mental health services should be rationally planned on the basis of evidence, efficient in their operation, and bring maximum benefit to patients using them. These goals are challenging, and require the use of a coherent theoretical framework when considering how best to meet them. A two-dimensional matrix model has been proposed by Thornicroft & Tansella (1999) for characterising mental health services. The model comprises a temporal dimension with input, process and outcome phases, and a geographical dimension with patient, local and country/regional levels.

In this model, the *inputs* are the resources which are put into the mental health care system. This is typically taken to include staff characteristics (*e.g.* number, training, morale), resource allocations (*e.g.* money, buildings, equipment) and invisible inputs (*e.g.* working relationships, laws, media representations of mental health). Some commentators also include patient characteristics as an input variable (Lyons *et al.*, 1997). The *processes* are the activities which take place to deliver mental health services. At the country/regional level processes will normally be expressed as rates of events (*e.g.* of compulsory detention) relative to the general population, at the local level as rates or levels relative to the catchment-area (*e.g.* annual treated prevalence), and at the patient level as the type, length and frequency of contact for individual patients (*e.g.* eight 1-hour sessions of cognitive therapy). Finally, *outcomes* are the visible effects of receiving mental health care, or *"the effect on a patient's health status attributable to an intervention by a health professional or health service'* (Andrews *et al.*, 1994, p. 3). Outcomes can also be assessed at each geographical level, ranging from the monitoring of national targets at country/regional level to evaluating the impact of mental health care for individual patients.

Three influences have led to a de-emphasis on outcomes, especially at the patient level. First, the focus of politicians, service funders and managers has tended to be on cost containment or reduction. This has led to the routine collection of data on inputs (such as relative and absolute funding levels) and processes (such as bed days), rather than outcomes. Second, outcome data at each level are very different. It is tempting to think that aggregating outcome data for individual patients will give information at the local or country/regional information. However, in practice it is difficult to aggregate information about individuals at the local level, and almost impossible to infer population outcomes solely using data about individual patients (Slade & Glover, 2000). This has led to the category error of using input and process measures as proxy variables for outcome (Jenkins *et al.*, 1994). Finally, even agreeing what constitutes a positive outcome is especially problematic in mental health (Slade, in press). Many groups will have an interest in the effects of mental health services, including people using and directly providing the service, managers, service funders, researchers, politicians and tax-payers. Each group will be interested in differing and sometimes incompatible outcomes. As examples, the patient who has fewer episodes of mania as a result of treatment may view this change as a negative outcome, and the tax-payer's desire for a safe community may result in increased compulsory detention rates.

Overall, outcomes in mental health have received less attention than inputs and processes, and outcomes at the patient level have received less attention than outcomes at the local and country/regional levels. Recently, this has been changing.

PATIENT-LEVEL OUTCOMES IN MENTAL HEALTH

North American behavioural healthcare services (including mental health services) have the longest history of using outcome data based on individual patient assessments (Lyons *et al.*, 1997). The history is worth briefly reviewing. In the United States, the costs of mental

health care rose from 4% in the early 1980s to nearly 25% by the early 1990s (Dickey & Azeni, 1992). This led to a focus on 'managed care', initially envisaged as a cost containment procedure and driven by pressure from health insurance companies. The first approach used *Diagnostic Related Groups* (DRGs), with each DRG being associated with access and benefit levels (i.e. limitations). This approach did not result in equity and a fair matching of need with resources, and so was superseded by an emphasis on outcomes rather than diagnosis. Several systems for routinely assessing outcome have been developed, such as the condition-specific Outcomes Module approach (Smith *et al.*, 1997a) and the COMPASS system which assesses predictors of outcome (Sperry *et al.*, 1996). The use of these systems is driven by the need to demonstrate value-for-money in the service, so they are normally used to generate aggregated data to provide 'report cards' on the service. Lyons *et al.* (1997) also give examples of baseline patient-level variables being used to give a projected rate of improvement, so that ongoing monitoring of outcome variables can inform decision-making about whether therapy should be continued (if the projected improvement matches observed progress), or altered or terminated (if it does not). Pressure to develop such systems can be expected to increase in European services, since they have the potential both to inform resource allocation decisions and to improve individual care.

In Europe, some individual services have also monitored and used patient-level outcome data. Well-known examples include the *South Verona Outcomes Project* in the north east of Italy (Ruggeri *et al.*, 1998) and the Traumatic Stress Unit in South London (Marks, 1998). However, the routine collection of outcome data is not widespread. Despite this, the monitoring and use of patient-level outcome data to inform the treatment of individual patients may be worth considering.

Conceptual basis

The North American experience led to the identification of four *sobering truths* about mental health care provision:

1. An unmanaged system has the potential for unlimited cost increases.

2. The supply-induced demand for mental health services has resulted from a dramatic increase in the number of mental health professionals.

3. Certain treatments, no matter how compelling, may not be worth their cost.

4. There is wide and inexplicable variation in the quality of the clinical and economic performance of individual providers. (Lyons *et al.*, 1997, p. 12)

These can by summarised in the axiom that *variation is the enemy of quality*. In response to these issues, a conceptual framework was put forward by Paul Ellwood in his 1988 Shattuck Lecture (Ellwood, 1988). He proposed a technology of *outcomes management*, comprising four elements:

(a) greater reliance on standards and guidelines for selecting appropriate interventions;

(b) the routine and systematic measurement of the functioning and well-being of patients, along with disease-specific clinical outcomes, at appropriate time intervals;

(c) the pooling of clinical and outcome data on a massive scale;

(d) the analysis and dissemination of results from the segment of the database most appropriate to the concerns of each decision-maker.

At the patient level, elements (b) and (d) suggest that clinical care would be improved if outcomes were assessed routinely with each patient, and then fed back into the care process to inform subsequent treatment and care – a process which will be called *routine outcome assessment*. Although used (if not highly evaluated) in North America, is there any evidence that routine outcome assessment is a cost-effective way of working in a European context?

Evidence for routine outcome assessment

Asking patients or staff to assess outcome leads to a beneficial focus on outcome in planning and evaluating treatment (Marks, 1998). For example, simply asking staff to rate whether the patient has a good social network of support may highlight the patient's social isolation, and prompt therapeutic intervention. This approach has been most developed in cognitive therapy, a psychological intervention which was developed for use initially in depression (*e.g.* Beck *et al.*, 1979), but has now been found to be effective in modified form with a wide range of psychiatric disorders (Roth & Fonagy, 1997). The technique of patient-based outcome monitoring involves the patient rating the target variable (*e.g.* level of depression) over time, with the results charted and shared with the therapist. This technique has many benefits: it sets up the expectation of change, it helps patients to feel heard, when the chart indicates improvement it reinforces change and challenges the common patient belief that they are making no progress, when the chart indicates dete-

rioration it contextualises this outcome as a 'blip' rather than an inexorable downward spiral, and finally it identifies whether the therapy is in fact working and prompts re-evaluation of the therapy when improvement is not evident.

More generally, deciding on what treatment to provide involves both deductive reasoning (starting with an idea, developing a theory and testable hypotheses, and then gathering data to confirm or contradict the hypotheses) and inductive reasoning (using observations to generate ideas and hypotheses, which are then tested by gathering further data) (Bowling, 1997). Most treatment planning by experienced clinicians is based on deductive reasoning – the application of a general theory (*e.g.* the conceptual framework provided by a diagnosis) to an individual patient. By contrast, routine outcome assessment may facilitate inductive reasoning, by prompting questions such as *"Why didn't treatment X work with patient Y? I wonder if it's because..."*. Such reflection has the potential to improve outcome, either directly (through changing the content of care) or as an effect modifier (by improving the process of care). This approach to treatment planning can be termed *reflective clinical practice*.

What is reflective clinical practice? The concept is most easily explained with reference to the experiential learning model proposed by Lewin and elaborated by Kolb (1984), shown in figure 1.

This model of learning and action can be applied to clinical settings. Busy routine clinical practice is characterised by many demands (experiences), which require processing by clinicians (reflection), matching with their theoretical basis acquired during training, and implementation (planned experimentation). Reflective clinical practice involves passing around this cycle in a conscious ('reflecting') manner. However, over time clinicians become skilled in these processes, and develop internal action plans based on a wide range of experiences. Using terminology from cognitive psychology, declarative (factual and academic) knowledge becomes transformed into procedural (skills-based) knowledge through practice (Anderson, 1983). Once procedural knowledge has been developed, problem solving requires accurate pattern recognition (diagnosis or formulation) to activate the appropriate procedural knowledge, stored as a problem-solving template. This approach can be termed *automated problem-solving*. Reflective clinical practice and automated problem-solving lie at opposite ends of a continuum, with automated problem-solving characterised by minimal conscious processing, and reflective clinical practice by maximal conscious processing (or reflection).

Research on problem-solving indicates that problems lie on a continuum from well-defined to ill-defined (Kahney, 1986). Well-defined problems are characterised by the availability of full information about the *initial state* of the problem, *the goal state*, the legal *operators* (things that can be done to solve the problem) and the *operator restrictions* (constraints on the application of operators). The less adequate the available information is, the more the problem becomes ill-defined. In the context of clinical decision-making, the initial state is the assessment, the goal state is the intended outcome of care, the operators

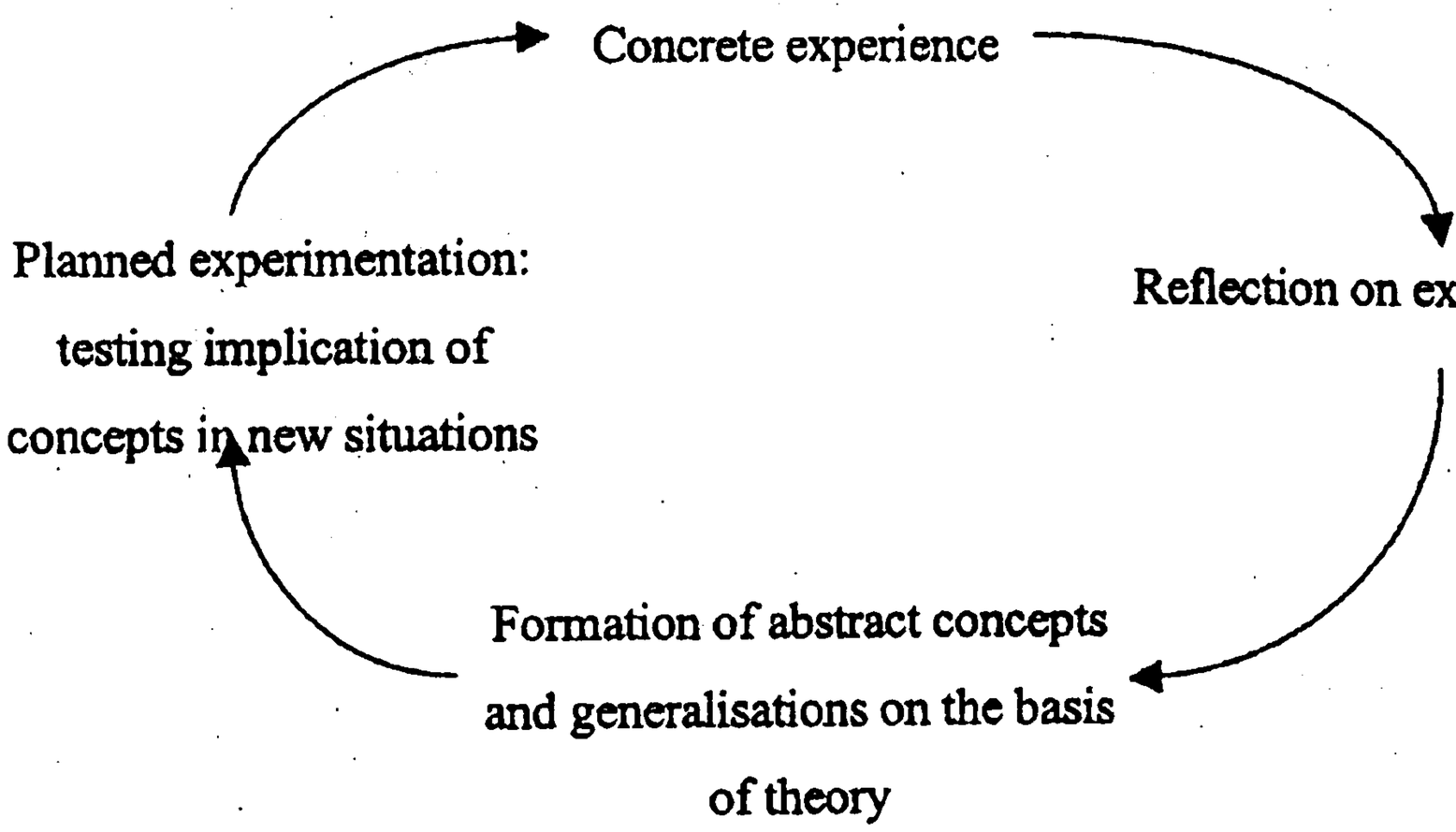


Figure 1 - *Lewin's Model of Experiential Learning.*

are potential treatments, and the operator restrictions are the constraints on the treatments which can be offered.

Automated problem-solving is most effective with well-defined problems – the approach can be rapidly applied, without time-consuming thought and reflection, to many well-defined problems. It is asserted here that clinical decision-making by experienced clinicians relies mainly on this approach. For example, the culture of mental health care is to value quantity of patients seen, with caseload often being used as a proxy for work effort. Sustaining a high caseload requires the frequent use of automated problem-solving, to ensure that defensible decisions are made as often as possible. There are limited incentives to reflect on experience – the nearest to this is when care plans are reviewed, but this is often in the context of a multidisciplinary review, rather than during ongoing clinical work. Thinking or reflection time can become seen as a dispensable luxury in meeting otherwise overwhelming clinical demands, with automated problem-solving becoming the dominant approach.

The difficulty with automated problem-solving is that it does not work well with ill-defined problems, which are exactly the type of problem which often arise in clinical practice. The patient may choose not to disclose some information, or the clinician may not ask. There may be a mismatch between patient and clinician goals for treatment. The apparent range of treatment options may be constrained by professional training. The constraints on the treatment that can be offered may be unclear. Ill-defined problems require reflection time to reach an optimal solution. This means that best quality care which maximises outcome requires more reflective clinical practice and less automated problem-solving. If mental health services are to maximise quality of care, then there needs to be active encouragement, facilitation and rewarding of reflective clinical practice in the work of individual clinicians. The focus on outcomes which results from routine outcome assessment will contribute to this goal.

An analogous process to reflective clinical practice can also be inferred for patients. They will have views about the causes of their own problems (initial state), what would constitute a successful outcome (goal state), what help (operators) is likely to be beneficial, and the relative importance of, for example, medication and side-effects (operator constraints). Patients clearly differ in the extent to which they want to be involved in decisions about their care, but preferred style may change over time – becoming more involved as they become empowered to express their views (Guadagnoli & Ward, 1998). Using routine outcome assessment will support the patient's ability to reflect on and discuss their treatment.

Routine outcome assessment – the realities

If routine outcome assessment has the potential to improve mental health care, then why is it not in widespread use? At least two reasons can be identified: there is a lack of suitable outcome measures, and clinicians remain unconvinced of the benefits.

Most outcome assessments have been developed using methodologies which lead to measures suitable for research use, and less is known about how to develop assessments intended for routine use. This characteristic of *feasibility* has been defined as a psychometric property of an outcome measure indicating *"the extent to which it is suitable for use on a routine, sustainable and meaningful basis in typical clinical settings, when used in a specified manner and for a specified purpose"* (Slade *et al.*, 1999a, p.245). Six characteristics of an assessment will enhance this property: being brief, simple, relevant, acceptable, available and valuable. Assessments have begun to be developed explicitly for routine use, so as to maximise feasibility, and an example of how to assess the feasibility of an outcome measure is given in (Slade *et al.*, 2001).

A substantial review relating to routine outcome assessment was published Andrews *et al.* (1994) in Australia. This review covered measures to assess symptoms, functioning, quality of life, burden and satisfaction, with inclusion criteria that a manual or published article describing the measure and its psychometric properties could be obtained. The resulting 95 identified measures were then evaluated in terms of their acceptability, applicability, practicality, reliability, validity and sensitivity to change. The five outcome measures which were identified as best meeting these criteria were the *Role Functioning Scale* (RFS) (McPheeters *et al.*, 1984), the *Behaviour and Symptom Identification Scale* – 32 item (BASIS-32) (Eisen *et al.*, 1994), the *Health of the Nation Outcome Scales* (HoNOS) (Wing *et al.*, 1998), the *Medical Outcomes Study Short Form* (SF36) (Ware & Sherbourne, 1992), and the *Mental Health Inventory* (MHI) (Veit & Ware, 1983). The *Life Skills Profile* (LSP) (Rosen *et al.*, 1989) would have met these criteria apart from the cost for use. These six measures were then field-tested in Australia (Stedman *et al.*, 1997), and HoNOS and MHI were identified as offering the greatest potential for widespread use. More recent measures which were designed specifically for routine clinical use include the *Helping Alliance Scale* (HAS) (Priebe & Gruyters, 1993), the *Manchester Short Assessment of Quality of Life* (MANSA) (Priebe *et al.*, 1999), the *Functional Analysis of Care Environment* (FACE) (Clifford, 1999)

to measure mental and physical state, daily functioning and social relationships, the *Camberwell Assessment of Need Short Appraisal Schedule* (CANSAS) (Slade *et al.*, 1999b, Italian Version: Ruggeri & Tansella, 2000), the *Carers and Users Expectations of Service – Users Version* (CUES) (Lelliott *et al.*, 2001), and the *Threshold Assessment Grid* (TAG) (Slade *et al.*, 2000) to measure the severity of mental health problems.

Despite these developments, European clinicians remain in general unconvinced of the benefits of routine outcome assessment, and (unlike North American colleagues) are not normally required to undertake any formal assessment of outcome. At the individual patient level, at least three steps are involved: data collection, management and feedback into the clinical decision-making process. Data collection does not necessarily have to involve clinical staff, since it can potentially be done in part by clinically untrained staff such as receptionists (Slade, in press). In practice, however, most approaches to routine outcome assessment do include staff-rated assessments. The entry, management, analysis and preparation of feedback of the data can in theory also be automated, but again the current level of information system still normally requires human involvement. Finally, the use of the feedback information will require clinician time, in integrating the supplied information with other clinical assessment information when planning care. Marks (1998) estimates that the use of even a rudimentary outcome measure adds about 10% to the time spent by the clinician on the patient's care. The best way to identify whether this is cost-effective is through evaluating the costs and benefits of routine outcome assessment. Two European studies evaluating routine outcome assessment are planned or currently underway: the European Union-funded MECCA Study at sites in Germany, Netherlands, Spain, Sweden, Switzerland, and England, and the UK Medical Research Council-funded FOCUS Study in South London.

What issues need to be considered by clinicians and services wanting to make outcomes more central to the planning of treatment?

Implementing routine outcome assessment

Several commentators have identified principles for implementing routine outcome assessment (*e.g.* Lyons *et al.*, 1997; Huxley, 1998; Trauer, 1998). The most widely-cited set of principles were developed by the Outcomes Roundtable, an organisation whose main joint sponsors are John Hopkins University and the National Alliance for the Mentally Ill (Smith *et al.*, 1997b). They identified 12 principles for routine outcome assessment:

1 *Outcomes assessments should be appropriate to the application or question being answered* – understanding the relationship between the mental health problem, the treatment and the resulting health status for individual patients requires disorder-specific assessment measures, whereas understanding this relationship for groups of patients requires generic assessment measures.
2 *Tools for assessing outcomes should have demonstrated validity and reliability and must be sensitive to clinically important change over time* – the reduction in data quality attributable to the assessment measure should be minimised.
3 *Outcomes assessments should always include the patient's perspective; outcomes assessments obtained from providers and family members may enhance what is learned* – especially in mental health, many aspects of the impact of treatment can only be assessed by the patient.
4 *Outcomes assessment systems should place minimal burden on the respondent and have the ability to be adapted to different health care systems* – the main focus of the treatment setting should be treatment, not measurement.
5 *Outcomes assessments should include general health status as well as mental health status* – physical health and social functioning mediate the relationship between mental health care and health gain.
6 *Outcomes assessments should include measures of the patient's evaluation of treatment and outcomes* – information about the patient's views on their treatment should inform the planning of their future care.
7 *Outcomes assessment tools should quantify the type and extent of treatment the patient receives (the process of care) for the target condition in order to understand the clinical relationship between the outcomes of care and treatment* – improving quality requires information about both process and outcome.
8 *Outcomes assessment tools should include generic and disorder-specific information that is predictive of expected patient outcomes; this prognostic information may include case mix and severity characteristics that are associated with choice or of success with treatment* – if the intention is to compare groups of patients, then it is necessary to know if the groups of patients are similar in terms of prognostic variables which predict outcome.
9 *Outcomes assessment should include areas of personal functioning affected by the condition* – not all do-

mains of life need to be assessed for all patients.

10 *Outcomes should be initially assessed and reassessed at clinically meaningful points in time given the course of the disorder* – assessment at hospital discharge is less meaningful (because most patients will have improved) than assessment at some fixed time after discharge.

11 *Outcomes assessments should use an appropriate scientific design and representative sample* – the most common error is drawing too strong a conclusion from data provided by a small sample, which is likely to be un-representative.

12 *Assessing outcomes of patients who leave treatment prematurely is as important as assessing outcomes of those who are still in treatment at the time of follow-up* – the best approach to identifying the impact of treatment is to consider all patients who received it.

To these twelve scientific principles can be added seven pragmatic principles based on experience in implementing routine outcome assessment in London:

1. *Identify the resourcing available, and then plan the level of implementation* – problems arise when an individual or a team start using an outcome measure with high enthusiasm, but no infrastructure for analysing and feeding back the resulting data. The pile of unanalysed data grows (literally), until the project is overtly abandoned or covertly sabotaged through disengagement and non-completion. The resulting disillusionment makes any subsequent attempt to introduce outcome assessment more difficult. It is preferable not to start, or to start on a smaller but sustainable scale (*e.g.* one team, instead of a service), than to start and then fail to use the information collected.
2. *Match the change to the culture* – start with a process measure (such as therapeutic alliance) in organisations emphasising process issues, or start with an outcome measure in organisations emphasising health gain. Try to move the culture towards being a learning organisation, which actively collects and uses information to improve productivity, quality, profitability and morale (Senge, 1990).
3. *Start small* – identify one key issue which has face validity, and then identify one question whose answer would inform decision-making about that issue, and then identify one simple way of getting relevant information to answer that question, and then develop one sustainable approach to analysing and feeding back the resulting data. The goal is to get enough information of adequate quality to start changing the culture towards valuing outcome information.
4. *Ensure visible and high-profile local ownership of this (or any) attempt to change working practices* – develop and support product champions: people of high local credibility who actively support the introduction of routine outcome assessment.
5. *The choice of outcome measure(s) should be one of the last decisions made* – decisions about what to measure, in what way, by whom and how often should all precede (and inform) decisions about which specific outcome measure(s) to use. The practice change which begins with the decision to use a particular scale will often fail through inadequate levels of local enthusiasm (apart from the change instigator) and resourcing.
6. *Get the assessment forms and processes and analysis and feedback right, first time* – several iterations of piloting and amending are likely to be necessary. If resources are not available for this piloting, then resources are not available to measure outcomes routinely.
7. *Distinguish ruthlessly between fact and speculation in the assessment data, and only feed back factual information* – the recipients of the outcome data will put their own interpretation on the data. Speculating about the meaning when feeding back results (*e.g.* by comparing teams when no input or process data are available) will undermine credibility.

In summary, it has been argued that the importance of routine outcome assessment will grow, both because of financial pressures and because it leads to improved clinical care. Clinicians have the opportunity to benefit from this change, if they can lead the incremental introduction of routine collection, analysis and use of outcome information.

REFERENCES

Anderson J.R. (1983). *The Architecture of Cognition.* Harvard University Press: Harvard.

Andrews G., Peters L. & Teeson M. (1994). *Measurement of Consumer Outcome in Mental Health: a Report to the National Mental Health Information Strategy Committee.* Clinical Research Unit for Anxiety Disorders: Sydney.

Beck A.T., Rush A.J., Shaw B.F. & Emery G. (1979). *Cognitive Therapy of Depression.* Guilford Press: New York.

Bowling A. (1997). *Research Methods in Health: Investigating Health and Health Services.* Open University Press: Milton Keynes.

Clifford P. (1999). The FACE Recording and Measurement System: a scientific approach to person-based information. *Bulletin of the Menninger Clinic* 63, 305-331.

Dickey B. & Azeni H. (1992). Data watch: impact of managed care on mental health services. *Health Affairs* 11, 197-204.

Eisen S.V., Dill D.L. & Grob M.C. (1994). Reliability and validity of a brief patient-report instrument for psychiatric outcome evalua-

tion. *Hospital and Community Psychiatry* 45, 242-247.

Ellwood P. (1988). Outcomes management: a technology of patient experience. *New England Journal of Medicine* 318, 1549-1556.

Guadagnoli E. & Ward P. (1998). Patient participation in decision-making. *Social Science and Medicine* 47, 329-339.

Huxley P. (1998). Outcomes management in mental health: a brief review. *Journal of Mental Health* 7, 273-283.

Jenkins R., Brooksbank D. & Miller E. (1994). Ageing in learning difficulties: the development of health care outcome indicators. *Journal of Intellectual Disability Research* 38, 257-264.

Kahney H. (1986). *Problem Solving: a Cognitive Approach.* Open University Press: Milton Keynes.

Kolb D.A. (1984). *Experiential Learning: Experience as the Source of Learning and Development.* Prentice-Hall: New Jersey.

Lelliott P., Beevor A., Hogman J., Hogman G., Hyslop J., Lathlean J. & Ward M. (2001). Carers' and Users' Expectations of Services – User version (CUES-U): a new instrument to measure the experience of users of mental health services. *British Journal of Psychiatry* 179, 67-72.

Lyons J.S., Howard K.I., O'Mahoney M.T. & Lish J.D. (1997). *The Measurement & Management of Clinical Outcomes in Mental Health.* John Wiley & Sons: New York.

Marks I. (1998). Overcoming obstacles to routine outcome measurement. *British Journal of Psychiatry* 173, 281-286.

McPheeters H.L. (1984). Statewide mental health outcome evaluation: a perspective of two southern states. *Community Mental Health Journal* 20, 44-55.

Priebe S. & Gruyters T. (1993). The role of the helping alliance in psychiatric community care: a prospective study. *Journal of Nervous and Mental Disease* 181, 552-557.

Priebe S., Huxley P., Knight S. & Evans S. (1999). Application and results of the Manchester Short Assessment of quality of life. *International Journal of Social Psychiatry* 45, 7-12.

Rosen A., Hadzi-Pavlovic D. & Parker G. (1989). The life skills profile: a measure assessing function and disability in schizophrenia. *Schizophrenia Bulletin* 15, 325-337.

Roth A. & Fonagy P. (1997). *What Works forWhom?* Guilford Press: New York.

Ruggeri M. & Tansella M. (2000). *CAN: Uno Strumento per la Valutazione dei Bisogni nei Soggetti con Disturbi Mentali Gravi.* CIC: Roma.

Ruggeri M., Biggeri A,. Rucci P. & Tansella M. (1998). Multivariate analysis of outcome of mental health care using graphical chain models. The South-Verona Outcome Project 1. *Psychological Medicine* 28, 1421-1431.

Senge P. (1990). *The Fifth Discipline: the Art and Practice of the Learning Organization.* Doubleday/Currency: New York.

Slade M. (in press). Routine outcome assessment in mental health services. *Psychological Medicine.*

Slade M. & Glover G. (2000). The needs of people with mental disorder. In *Textbook of Community Psychiatry* (ed. G. Thornicroft and G. Szmukler), pp. 117-127. Oxford University Press: Oxford.

Slade M., Thornicroft G. & Glover G. (1999a). The feasibility of routine outcome measures in mental health. *Social Psychiatry and Psychiatric Epidemiology* 34, 243-249.

Slade M., Loftus L., Phelan M., Thornicroft G. & Wykes T. (1999b). *The Camberwell Assessment of Need.* Gaskell: London.

Slade M., Powell R., Rosen A. & Strathdee G. (2000). Threshold Assessment Grid (TAG): the development of a valid and brief scale to assess the severity of mental illness. *Social Psychiatry and Psychiatric Epidemiology* 35, 78-85.

Slade M., Cahill C., Kelsey W., Powell R., Strathdee G. & Valiakalayil A. (2001). Threshold 3: The feasibility of the Threshold Assessment Grid (TAG) for routine assessment of the severity of mental health problems. *Social Psychiatry and Psychiatric Epidemiology* 36, 516-521.

Smith G.R., Rost K., Fischer E., Burnam M. & Burns B. (1997a). Implementing outcomes management systems in mental health settings. *Psychiatric Services* 48, 364-368.

Smith G.R., Manderscheid R., Flynn L. & Steinwachs D. (1997b). Principles for assessment of patient outcomes in mental health care. *Psychiatric Services* 48, 1033-1036.

Sperry L., Brill P., Howard K.I. & Grissom G. (1996). *Treatment Outcomes in Psychotherapy and Psychiatric Interventions.* Brunner/Mazel: New York.

Stedman T., Yellowlees P., Mellsop G., Clarke R. & Drake S. (1997). *Measuring Consumer Outcomes in Mental Health.* Department of Health and Family Services: Canberra, ACT.

Thornicroft G. & Tansella M. (1999). *The Mental Health Matrix. A Manual to Improve Services.* Cambridge University Press: Cambridge.

Trauer T. (1998). Issues in the assessment of outcome in mental health. *Australian and New Zealand Journal of Psychiatry* 32, 337-343.

Veit C.T. & Ware J.E. (1983). The structure of psychological distress and well-being in general populations. *Journal of Consulting & Clinical Psychology* 51, 730-742.

Ware J.E. & Sherbourne C.D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care* 30, 473-483.

Wing J.K., Beevor A.S., Curtis R.H., Park S.B., Hadden S. & Burns A. (1998). Health of the Nation Outcome Scales (HoNOS) - Research and development. *British Journal of Psychiatry* 172, 11-18.

Psychological Medicine, 2002, **32**, 1339–1343. © 2002 Cambridge University Press
DOI: 10.1017/S0033291701004974 Printed in the United Kingdom

EDITORIAL

Routine outcome assessment in mental health services[1]

Measuring and interpreting outcome is more difficult in mental health services than in some other areas of health care, for at least five reasons. First, the effect of the treatment may be to slow decline or to maintain the current level, so the score on the outcome measure itself may not improve (or may even get worse) despite best quality clinical care. Secondly, the best available evidence in the United Kingdom indicates that clinical and social variables predict no more than 30% of the variance in an individual's quality of life (UK700 Group, 1999). Thirdly, different types of outcome are desynchronous (e.g. Drury *et al.* 1996), changing at different rates during an intervention. Fourthly, there may not be agreement regarding what is a positive change in outcome – the patient who has fewer episodes of mania as a result of treatment may see this as a negative outcome. Finally, three levels of mental health service can be differentiated: treatment (specific interventions),'programme (combination of different treatment components); and system (all programmes for a defined target group in a given area) (Burns & Priebe, 1996). The outcome data needed to evaluate each level will be very different.

The solution to these issues that has evolved in research studies has been to assess a wide range of treatment and programme-level outcomes, from multiple perspectives. For example, the programme-level PRiSM Psychosis Study evaluated two models of community care for people with psychotic diagnoses (Thornicroft *et al.* 1998). The outcome domains assessed by interviewing the patient were symptomatology, needs, quality of life, services being received (to allow economic analysis), social network and satisfaction with care. The outcome domains assessed by interviewing staff were global level of functioning, needs and social behaviour, and by interviewing carers were their experience of care-giving and their own symptomatology. All interviews were conducted by researchers. In general, most research and evaluation studies take place in 'research contexts where specifically funded and trained external raters parachute into routine clinical settings in order to guarantee the validity and reliability of study measures' (Harrison & Eaton, 1999, p. 187).

RESEARCH FINDINGS

This approach has led to the identification of several consistent findings. The most important outcome is quality of life, and the best predictor of quality of life is level of unmet need, which is a better predictor than diagnosis, symptomatology or other social or clinical variables (McCrone & Strathdee, 1994; Slade *et al.* 1999*a*; UK700 Group, 1999). Staff and patient assessments of need differ, so both should be considered (Sainfort *et al.* 1996; Slade *et al.* 1998; Hansson *et al.* 2001). In this issue Gilbody *et al.* (2002) review studies of routine administration of quality of life and needs assessment instruments. More generally, outcome should be considered multi-dimensionally, by measuring multiple outcome domains (Biggeri *et al.* 1996; Clifford, 1998). Process variables are also important. There is a moderate but reliable relationship between alliance as rated by either staff or patient and outcome, and this relationship is not influenced by other moderator variables such as outcome measure or rater, time or type of alliance assessment, type of treatment or publication status of study (Keijsers *et al.* 2000; Martin *et al.* 2000). There is also a robust relationship between collaboration in the relationship and medication compliance (Fenton *et al.* 1997).

[1] Address for correspondence: Dr Mike Slade, Health Services Research Department, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF.

In summary, mental health outcomes research indicates that care should be provided on the basis of need so as to improve quality of life, that attention should be paid to the therapeutic relationship, and that outcome should be considered from more than one perspective. This involves assessing these outcomes in some way. Are outcomes such as quality of life, needs and therapeutic alliance being routinely measured in practice? Examples from the United States, England and Australia will be considered.

CURRENT PRACTICE

In the United States, the rising costs of mental health care resulted in the rapid introduction of managed care in the 1980s (Dickey & Azeni, 1992). Initially envisaged as a cost containment procedure and driven to some extent by pressure from health insurance companies, the approach was based on the identification of diagnostic related groups (DRGs), with each DRG being associated with access and benefit levels (i.e. limitations). The failure of the DRG system to result in equity and a fair matching of need with resources has led to this approach being superseded by an emphasis on outcomes rather than diagnosis. However, consideration of needs remains entirely absent, for example not warranting a section in the 820-page *Handbook of Psychiatric Measures* by the American Psychiatric Association (2000).

In England, active commissioning of mental health services was introduced over the same period, driven by the political perspective that systems work most efficiently when they follow the commercial model of rewarding the matching of supply and demand. The structure of commissioning changed during the decade, moving from District Health Authorities, through to a combination of Health Authority and fund-holding GPs, to the development of Primary Care Groups and Primary Care Trusts. However, the common theme was that service commissioners wanted to ensure that they got what they paid for. Since commissioners pay for structure (e.g. hospitals, mental health staff) and processes (e.g. Mental Health Act assessment, therapy), routinely collected data focused on these aspects. The result was that most if not all data collected for local (commissioner) or national return focused on aspects such as where, when, how often and for how long patients were seen, either in individual meetings or over extended periods ('consultant episodes'). In the late 1990s, a new emphasis on quality and outcome began to emerge.

In Australia, by contrast, the process of consumer involvement in mental health services is more advanced, and substantial efforts have been directed towards considering how outcomes should be monitored at the level of the individual patient. A seminal report by Andrews and colleagues (1994) identified specific outcome measures that should be considered for use in Australian mental health services. These measures were field-tested by an independent research team, resulting in specific recommendations for services (Stedman *et al.* 1997), with some assessments now used routinely throughout many adult mental health services.

International attempts to assess outcome have focused on informing management decision-making at the system level, with an emphasis on the perceived benefits of system-level as opposed to treatment-level outcome indicators.

ROUTINE OUTCOME ASSESSMENT

The ongoing measurement of treatment-level outcomes in routine mental health service – which will be referred to as routine outcome assessment – could be justified for both ethical and scientific reasons. Ethically, it is important to ensure that the treatment being provided in routine services is of the highest quality, which can only be done by monitoring its impact. Scientifically, although a fair amount is known about the efficacy of a range of treatments (established in research studies), far less is known about the effectiveness or cost-effectiveness of treatments when used in routine mental health services.

It is, therefore, worth considering why routine outcome assessment is currently not undertaken in most mental health services. A number of possible reasons for this have been suggested, including

lack of appropriate instruments, time and incentives (financial and professional) to offset the costs of monitoring outcome, and lack of expectations from senior staff that junior staff collect outcome data (Huxley, 1998; Marks, 1998; Slade *et al.* 1999*b*; Walter *et al.* 1998). Purely in terms of the assessment process, there no consensus regarding what outcome domains to include, who to ask when assessing, and what assessment measures to use (Clifford, 1998). Although these difficulties can be addressed, as shown by the Australian developments, clinicians remain in general unconvinced of the benefits of routinely monitoring outcome. The remainder of this editorial will argue that routine outcome assessment has the potential to be of benefit at the treatment level.

BENEFITS FOR PATIENTS

Emerging evidence suggests that routine outcome assessment has the potential to inform the treatment of individual patients. There is strong evidence that routine outcome assessment results in a beneficial focus on outcome in evaluating treatment approaches (Biggeri *et al.* 1996; Marks, 1998). Furthermore, there is emerging evidence from trials of cognitive therapy, in which the results of patient-based outcome monitoring are charted over time and shared with the therapist. The rationale for this monitoring is to set up the expectation of change, to reality test the common patient belief that they are making no progress, and to identify if indeed the therapy is working. Many patients benefit from this technique, both in 'feeling heard' and contextualizing the outcome as a 'blip' rather than a downward spiral when the chart indicates deterioration, and in reinforcing change when the chart indicates improvement. From a therapist perspective, the charting of patient-based outcomes can act as a motivator to re-evaluate the treatment plan where no improvement is evident. Such reflection has the potential to improve outcome, either directly (through changing the content of care) or as an effect modifier (by improving process issues).

How can reflective practice be facilitated? Simple feedback to the staff may be insufficient (Simon *et al.* 2000). One approach would be the routine collection of outcome data from both staff and patient perspectives, and then the routine feedback of these data to staff and patients. Either or both of completing the assessments and receiving the feedback may prompt reflection on the process or the content of care. The potential mechanisms of change for staff involved in such an approach are shown as a testable model in Fig. 1. Equivalent processes can be hypothesized for patients.

IMPLEMENTATION STRATEGIES

How should services wishing to implement routine outcome assessment proceed? The approach of measuring every plausible outcome from each relevant perspective cannot be directly transferred for at least three reasons. First, it requires the use of resources (e.g. interviewer time) which, while possible for efficacy studies, are unlikely to be available in routine services. Monitoring even a small number of outcome domains in routine practice is time-consuming – Marks (1998) estimates an extra 10% of the clinician's time is involved. Secondly, it entails duplication of effort (when two outcome measures co-vary to the extent that one is a fair proxy for another) and it can be wasteful of effort, either when data are collected but not analysed, or when data are collected and analysed but do not inform future treatment. It may be acceptable to absorb the adverse effects of duplication and waste of effort in research programmes, but in already over-stretched routine mental health services this is less possible. Thirdly, the measures themselves may not be 'feasible', or practical for use in routine clinical settings (Slade *et al.* 1999*b*).

Several recommendations flow from these observations. To support the implementation of routine outcome assessment, there is a need to identify the key outcomes to measure in mental health services – unmet need and quality of life have been proposed in this editorial, with therapeutic alliance as an important effect modifier. The development of assessment measures designed specifically for routine clinical use should be prioritized – there should be a moratorium on the development of further research-based assessments. Dedicated interviewers are not available

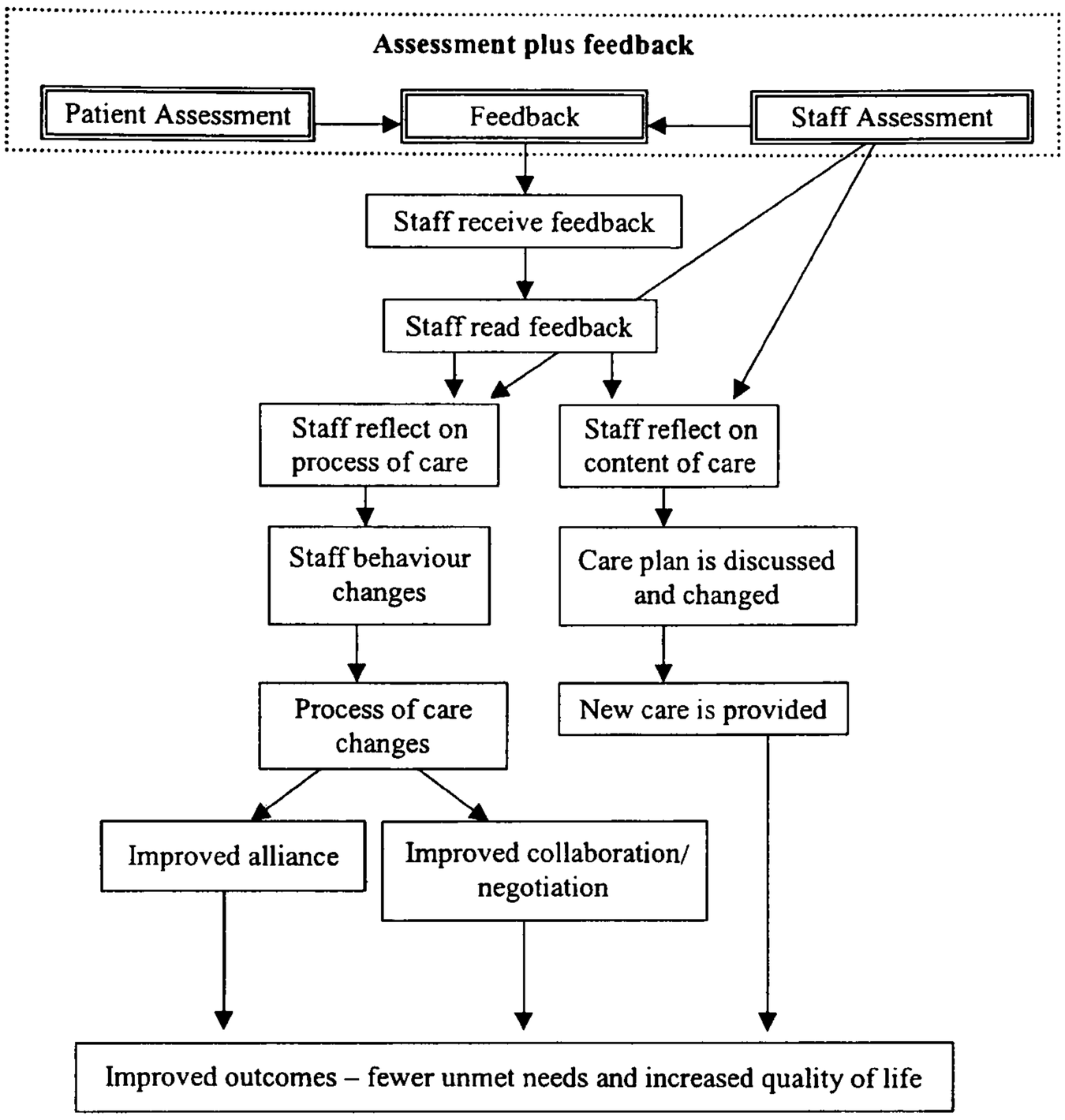


FIG. 1. Mechanisms of action for staff involved in routine outcome assessment and feedback.

in routine settings, so approaches need to be developed to collect outcome data for minimal cost in time and effort to staff and patients – the use of postal questionnaires, the internet and reception staff should all be explored. Audit Departments, which audit structure and process, should be replaced with Outcome Departments to support routine outcome assessment and feedback. Increasing access to and training in information technology should be a priority for all mental health staff, together with the development of an information infrastructure which supports clinical data collection and feedback – in England, the Mental Health Minimum Data Set work would be the obvious platform for this development. Finally, if quality of care is to be maximized, then there needs to be less of a focus on quantity (exemplified by caseload size) and more on promoting and rewarding reflective practice, both at the treatment and the programme level.

CONCLUSION

In the absence of sound empirical evidence about the relative costs and benefits, there is a risk of ill-conceived and haphazard attempts at routine outcome assessment, which will consume valuable resources, such as staff and patient time, for no evident benefit to the patient. There is some evidence of this happening already – Benjamin and colleagues (1995) reflect that the expectations of American policy-makers and service commissioners are that assessment will 'not be "too expensive", not show that the most expensive therapy is best, be easily comprehensible, address the things patients consider important, and, most importantly, save money' (p. 305).

Routine outcome assessment may in itself be of benefit for informing the treatment of patients. Current policy initiatives, the limited evidence base, and the recent development of standardized

outcome measures specifically intended for routine clinical use (e.g. Wing *et al.* 1998; Slade *et al.* 2000) all mean that studies to identify the costs and benefits for patient care of routine outcome assessment are urgently needed.

MIKE SLADE

REFERENCES

American Psychiatric Association (2000). *Handbook of Psychiatric Measures*. APA: Washington, DC.

Andrews, G., Peters, L. & Teeson, M. (1994). *Measurement of Consumer Outcome in Mental Health: A Report to the National Mental Health Information Strategy Committee*. Clinical Research Unit for Anxiety Disorders: Sydney.

Benjamin, K., Perfetto, E. & Greene, J. (1995). Public policy and the application of outcomes assessments: paradigms versus politics. *Medical Care* **33**, (Suppl.), AS299–305.

Biggeri, A., Rucci, P., Ruggeri, M. & Tansella, M. (1996). Multidimensional assessment of outcome: the analysis of conditional independence as an integrated statistical tool to model the relationships between variables. In *Mental Health Outcome Measures* (ed. G. Thornicroft G and M. Tansella), pp. 207–216. Springer: Berlin.

Burns, T. & Priebe, S. (1996). Mental health care systems and their characteristics: a proposal. *Acta Psychiatrica Scandinavica* **94**, 381–385.

Clifford, P. (1998). M is for outcome: the CORE outcomes initiative. *Journal of Mental Health* **7**, 19–24.

Dickey, B. & Azeni, H. (1992). Data watch: impact of managed care on mental health services. *Health Affairs* **11**, 197–204.

Drury, V., Birchwood, M., Cochrane, R. & Macmillan, F. (1996). Cognitive therapy and recovery from acute psychosis: a controlled trial. *British Journal of Psychiatry* **169**, 602–607.

Fenton, W. S., Blyler, C. R. & Heinssen, R. K. (1997). Determinants of medication compliance in schizophrenia. *Schizophrenia Bulletin* **23**, 637–651.

Gilbody, S. M., House, A. O. & Sheldon, T. (2002). Routine administration of Health Related Quality of Life (HRQoL) and needs assessment instruments to improve psychological outcome a systematic review. *Psychological Medicine* **32**, 1345–1356.

Hansson, L., Vinding, H. R., Mackeprang, T., Sourander, A., Werdelin, G., Bengtsson-Tops, A., Bjarnason, O., Dybbro, J., Nilsson, L., Sandlund, M., Sorgaard, K. & Middelboe, T. (2001). Comparison of key worker and patient assessment of needs in schizophrenic patients living in the community: a Nordic multicentre study. *Acta Psychiatrica Scandinavica* **103**, 45–51.

Harrison, G. & Eaton, W. (1999). From research world to real world. *Current Opinion in psychiatry* **12**, 187–189.

Huxley, P. (1998). Outcomes management in mental health: a brief review. *Journal of Mental Health* **7**, 273–283.

Keijsers, G. P., Schaap, C. P. & Hoogduin, C. A. (2000). The impact of interpersonal patient and therapist behavior on outcome in cognitive-behavioural therapy. A review of empirical studies. *Behavior Modification* **24**, 264–297.

McCrone, P. & Strathdee, G. (1994). Needs not diagnosis: towards a more rational approach to community mental health resourcing in Britain. *International Journal of Social Psychiatry* **40**, 79–86.

Marks, I. (1998). Overcoming obstacles to routine outcome measurement. *British Journal of Psychiatry* **173**, 281–6.

Martin, D. J., Garske, J. P. & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: a meta-analytic review. *Journal of Consulting and Clinical Psychology* **68**, 438–450.

Sainfort, F., Becker, M. & Diamond, R. (1996). Judgments of quality of life of individuals with severe mental disorders: patient self-report versus provider perspectives. *American Journal of Psychiatry* **153**, 497–502.

Simon, G. E., VonKorff, M., Rutter, C. & Wagner, E. (2000). Randomised trial of monitoring, feedback, and management of care by telephone to improve treatment of depression in primary care. *British Medical Journal* **320**, 550–444.

Slade, M., Phelan, M. & Thornicroft, G. (1998). A comparison of needs assessed by staff and an epidemiologically representative sample of patients with psychosis. *Psychological Medicine* **28**, 543–50.

Slade, M., Leese, M., Taylor, R. & Thornicroft, G. (1999*a*). The association between needs and quality of life in an epidemiologically representative sample of people with psychosis. *Acta Psychiatrica Scandinavica* **100**, 149–157.

Slade, M., Thornicroft, G. & Glover, G. (1999*b*). The feasibility of routine outcome measures in mental health. *Social Psychiatry and Psychiatric Epidemiology* **34**, 243–249.

Slade, M., Powell, R., Rosen, A. & Strathdee, G. (2000). Threshold Assessment Grid (TAG): the development of a valid and brief scale to assess the severity of mental illness. *Social Psychiatry and Psychiatric Epidemiology* **35**, 78–85.

Stedman, T., Yellowlees, P., Mellsop, G., Clarke, R. & Drake, S. (1997). *Measuring Consumer Outcomes in Mental Health*. Department of Health and Family Services: Canberra, ACT.

Thornicroft, G., Strathdee, G., Phelan, M., Holloway, F., Wykes, T., Dunn, G., McCrone, P., Leese, M., Johnson, S. & Szmukler, G. (1998). PRiSM Psychosis Study: rationale and design. *British Journal of psychiatry* **173**, 363–370.

UK700 Group. (1999). Predictors of quality of life in people with severe mental illness. *British Journal of Psychiatry* **175**, 426–432.

Walter, G., Cleary, N. & Rey, J. (1998). Attitudes of mental health personnel towards rating outcome. *Journal of Quality in Clinical Practice* **18**, 109–115.

Wing, J., Beevor, A., Curtis, R., Park, S., Hadden, S. & Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS. *British Journal of Psychiatry* **172**, 11–18.